The RNA-binding protein RBP10 controls a regulatory cascade that defines bloodstream-form trypanosome identity.

Elisha Mugo and Christine Clayton

DKFZ-ZMBH Alliance, Zentrum für Molekulare Biologie der Universität Heidelberg, University of Heidelberg. Im Neuenheimer Feld 282, D69120 Heidelberg, Germany.

Abstract

Gene expression control in the pathogen *Trypanosoma brucei* relies almost exclusively on post-transcriptional mechanisms, so RNA binding proteins must assume the burden that is usually borne by transcription factors. *T. brucei* multiply in the blood of mammals as bloodstream forms, and in the midgut of Tsetse flies as procyclic forms. We show here that a single RNA-binding protein, RBP10, defines the bloodstream-form trypanosome differentiation state. Depletion of RBP10 from bloodstream-form trypanosomes gives cells that can grow only as procyclic forms; conversely, expression of RBP10 in procyclic forms converts them to bloodstream forms. RBP10 binds to procyclic-specific mRNAs containing an UAUUUUUU motif, targeting them for translation repression and destruction. Products of RBP10 target mRNAs include not only the major procyclic surface protein and enzymes of energy metabolism, but also protein kinases and stage-specific RNA-binding proteins: consequently, alterations in RBP10 trigger a regulatory cascade.

Introduction

In multicellular organisms, cell differentiation is usually unidirectional, with stabilisation of differentiated states by epigenetic mechanisms, and reversal only in when healthy gene regulation is disrupted. In contrast, differentiation in unicellular organisms can be reversible, or involve cycles in which the direction of each transition is fixed. In each case differentiation is initiated by environmental stimuli, which activate signal transduction cascades; these in turn result in post-translational modifications that modify transcription factor activity. The consequent changes in mRNA synthesis, and hence protein synthesis, trigger a cascade of downstream regulatory events, affecting additional signalling molecules, transcription factors, and post-transcriptional modulators. Differentiation states are then stabilised by modification of DNA and chromatin.

Kinetoplastids are unicellular flagellates, and include many pathogens of mammals and plants, including the leishmanias and trypanosomes that cause human disease. Their life cycles involve a series of unidirectional transitions, which are driven by environmental changes as they move between mammalian and invertebrate hosts. The kinetoplastid studied in this paper, *Trypanosoma brucei*, causes human sleeping sickness in Africa and diseases of ruminants throughout the tropics. *T. brucei* multiply as bloodstream form trypomastigotes in the blood and tissue fluids of mammals, escaping immunity by antigenic variation of Variant Surface Glycoprotein (VSG) [1].

High cell density triggers growth arrest and differentiation to stumpy forms [2], which, upon uptake by a Tsetse fly, convert to proliferative procyclic forms in the midgut. Procyclic forms express EP and GPEET procyclin surface proteins [3], and have numerous metabolic adaptations including increased mitochondrial metabolism. In the following 2 weeks, trypanosomes migrate towards the salivary glands [4]. Epimastigotes, expressing the surface protein BARP [5], are found attached to the proventriculus and salivary glands; here, a subset of parasites undergoes meiosis and mating [6,7]. Finally, growth-arrested, VSG-expressing metacyclic forms are seen swimming freely in the salivary glands; these are now ready for infection of a new mammal. Trypanosome developmental transitions are marked by many changes in mRNA [8-12] and protein [13-16] abundance.

Despite extensive changes in gene expression through the *T. brucei* life cycle, and the existence of stable differentiation states, the normal paradigm for regulation does not apply. This is because - as in other Kinetoplastid protists - transcription by *T. brucei* RNA polymerase II is polycistronic, with no regulation at the level of individual mRNAs. Monocistronic mRNAs are excised by 5' *trans* splicing of a 39nt capped leader sequence, and by 3' polyadenylation [17,18]. Regulated is thus exclusively at the levels of mRNA processing, degradation, and translation [10-12,19-22]. The sequences required for regulation of translation and mRNA decay often lie in the 3'-untranslated regions (3'-UTRs) of the mRNAs. Many groups of trypanosome mRNAs are coordinately regulated [23] and in a few cases, co-regulated genes have been shown to share regulatory elements that are bound by specific RNA-binding proteins [24,25]. RNA-binding proteins thus assume the burden that is carried by transcription factors in other eukaryotes, by regulating distinct mRNA subsets [19,26]. Although an RNAi machinery exists, it is not essential for cell proliferation or differentiation [27], and microRNAs have not been found [28].

The very abundant *VSG* and procyclin mRNAs are made, exceptionally, by RNA polymerase I [29], and their transcription is developmentally regulated by epigenetic mechanisms [1]. Procyclin is not subject to antigenic variation, and the presence of even a small amount on the bloodstream-form surface would enable the development of host immunity. Suppression of procyclin expression is therefore essential for trypanosome pathogenicity. However, procyclin gene transcription is only 10-fold less active in bloodstream forms than in procyclic forms [30]. This means that additional suppressive mechanisms are required: the mRNAs are both extremely rapidly degraded and very poorly translated [31-35], and trafficking of procyclin proteins to the surface is suppressed [36]. A 26nt motif within the 3'-UTR of procyclin

mRNAs is required for their degradation and translational suppression [33,34]. However, until now, the protein(s) responsible for preventing procyclin expression were unknown.

RBP10 is a cytosolic protein with a single RNA recognition motif. Proteome analyses [16] and Western blot results [37] indicate that RBP10 is expressed in proliferating bloodstream forms, but not in stumpy forms or procyclics; *RBP10* mRNA also has low abundance in procyclic forms and in salivary gland trypanosomes [38]. Depletion of RBP10 by RNA interference (RNAi) in the bloodstream form is lethal, and forced RBP10 expression in procyclics inhibits growth [37]. Microarray analyses on cells with RBP10 RNAi or forced expression suggested a function for RBP10 either in stabilizing bloodstream-form specific mRNAs, or destabilizing procyclic-specific mRNAs, but it was unclear which effects were due to growth inhibition and the direct RNA targets of RBP10 were unknown.

In this paper we set out to identify the direct mRNA targets of RBP10 and to determine its role in mRNA control. We discovered that the mRNAs that are bound to RBP10 in bloodstream forms are preferentially targeted for translation inhibition and destruction, and that cells without RBP10 can survive only by differentiating to the procyclic form. Conversely, cells expressing RBP10 can survive only as bloodstream forms. The presence of RBP10 thus defines bloodstream-form trypanosome identity.

Results

Attachment of the RBP10 C-terminal domain to a reporter mRNA causes translation inhibition and mRNA degradation

Inducible expression of lambdaN-peptide-tagged RBP10 in cells containing a *CAT-boxB* reporter will result in attachment of the RBP10 to the reporter RNA via the lambdaN-boxB interaction. We previously showed that after 24h lambdaN-RBP10 induction, there was a decrease in boxB-tagged chloramphenicol acetyltransferase (*CAT*) reporter expression [37,39,40]. New deletion analyses (Fig 1) revealed that the C-terminal 88 amino acids of RBP10 was sufficient for the repressive effect (Fig 1 A, fragment F3). To investigate the mechanism in more detail we examined whether the *CAT-boxB* mRNA comigrated with polysomes in sucrose gradients. In the absence of tethered protein, the *CAT-boxB* reporter migrated in the dense sucrose fraction with large polysomes (Fig 1 B). With lambdaN-RBP10-F3 expression, *CAT-boxB* mRNA was reduced in abundance and the remainder migrated towards the top of the gradient. The control mRNA (encoding alpha-tubulin) was unaffected (Fig 1 B). The

RBP10 C-terminus thus prevents translation initiation and promotes mRNA destruction. Interestingly, although a homologue of RBP10 is present in all sequenced Kinetoplastids, only the RRM domain is conserved: this C-terminal sequence is specific to *Trypanosoma* (S1 Fig).

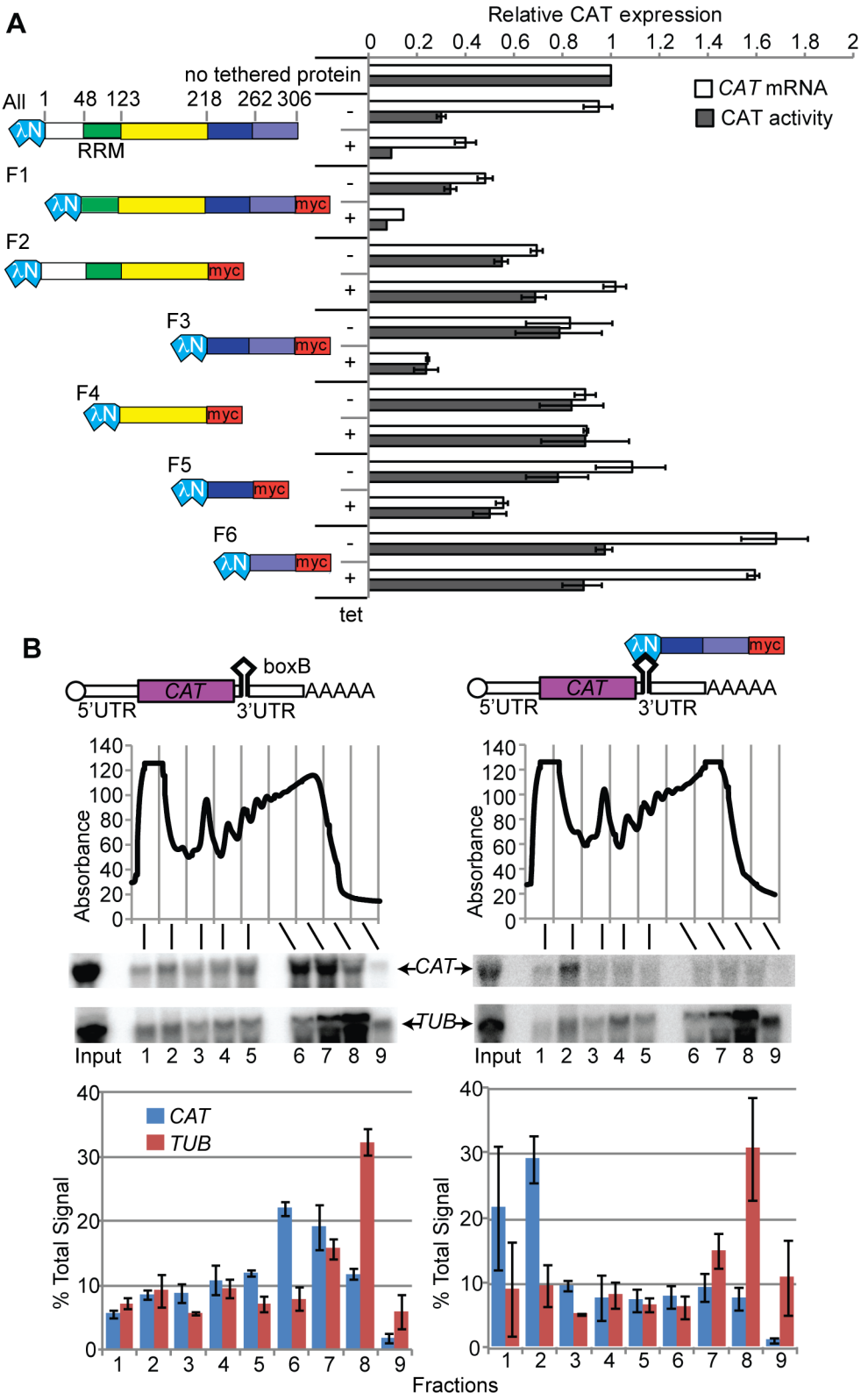


Fig 1

Tethering of the C-terminal portion of RBP10 to a reporter mRNA inhibits its translation and causes its destruction.

A. Various fragments of RBP10 were inducibly expressed with an N-terminal lambdaN peptide and a C-terminal myc tag, in a cell line expressing a *CAT* reporter mRNA with 5 boxB sequences in the 3'-UTR. Cartoons showing the fragments are on the left. The numbers above the full-length protein are amino acid residues. The effects on the amounts *CAT* mRNA and CAT protein were measured by Northern blot and enzyme assay, respectively. Levels are expressed as mean ± standard deviation for at least 3 replicates, relative to a line with no lambdaN peptide protein (top bars)

B. Bloodstream-form trypanosomes expressing the reporter in (A), and with tetracycline-inducible expression of lamdaN-RBP10-F3, were used. The reporter with or without tethered protein is shown at the top. Lysates from cells grown without (left) or with (right) tetracycline were subjected to sucrose gradient centrifugation. The panel below the cartoons show the absorption profiles at 254 nm as the fractions were collected. The migration of *CAT* and beta-tubulin (*TUB*) mRNAs on the gradient was detected by Northern blotting; representative blots are shown. The graphs show Northern signals, expressed as the percentage of the total signal, with arithmetic mean and standard deviation for three independent biological replicates.

RBP10 can interact with an eIF4E interacting protein

We investigated the mechanism of RBP10-mediated repression by looking for interaction partners. We made a bloodstream-form cell line containing a single copy of RBP10 bearing a cleavable N-terminal affinity purification tag (S2 Fig A). After tandem affinity purification the most consistent co-purifying proteins were components of the proteasome (S1 Table). While this could be a mechanism for repression, it is also possible that some of the affinity-tagged RBP10 was targeted to the proteasome because it was incorrectly folded. The only interaction partner linked to mRNA degradation was CAF40, a component of the NOT deadenylase complex, but since no other components of the complex were present the significance of the interaction was unclear. Furthermore, V5-CAF40 pull-down failed to identify any RBP10 peptides (data not shown). To find out whether other proteins - including those of unknown function - might be important we investigated direct interactions of RBP10 by 2-hybrid screening. Nearly 80 potential partners were seen (S2 Table), of

which 40 had not been seen in other screens. This suggested that RBP10 is prone to non-specific interactions. Potential partners included RBP10 itself, and four other RNA binding proteins. Interestingly, 8 of the 33 specifically interacting proteins had previously been show to suppress expression in the tethering assay [39,40]. There was however no overlap between the two-hybrid results and those from affinity purification.

A single fragment of a trypanosome eIF4E-interacting protein (4EIP) was scored as a possible RBP10 partner in the two-hybrid screen. Leishmania 4EIP interacts with eIF4E1 [41], which is one of the six trypanosomatid homologues of the cap-binding translation initiation factor eIF4E. In *T. brucei*, both 4EIP and eIF4E1 are extremely strong suppressors of reporter expression in the tethering assay [39,40]. We confirmed the two-hybrid interaction between full length RBP10 and 4EIP (Fig 2 A); and found that the interaction was mediated by the C-terminal 88 amino acids (F3) that are active in the tethering assay. This would be consistent with RBP10 acting via 4EIP recruitment. To test the interaction *in vivo*, we used trypanosomes expressing RBP10 that was N-terminally V5-tagged at the endogenous locus, then expressed C-terminally myc-tagged 4EIP from a strong promoter. Immunoprecipitation of myc-tagged 4EIP indeed pulled down a very small proportion of V5-RBP10 (Fig 2 B). It is therefore possible that RBP10 acts by recruiting 4EIP. However, depletion of 4EIP does not reproducibly inhibit bloodstream-form trypanosome growth ([42] and our preliminary results), which argues against this hypothesis.

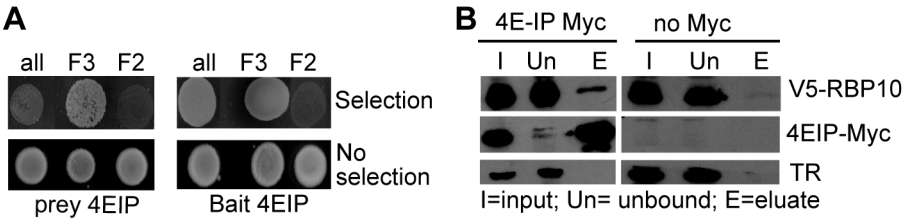


Fig 2

Interaction of RBP10 with 4E-IP

A. Two-hybrid assay showing yeast growth. Upper panels show selection against non-interacting pairs.

B. Immunoprecipitation of 4E-IP-Myc pulls down V5-tagged RBP10. The number of cell equivalents loaded for the eluate was 33x that of input and unbound fractions. The inverse precipitation did not confirm the interaction, but this could be because of an imbalance of protein abundances in the cells. (The abundance of 4E-IP is not known).

183

184 **RBP10 represses procyclic-form specific mRNAs**

185 To analyse the role of RBP10 in gene expression in bloodstream forms, we depleted
186 RBP10 by RNAi for 15h (S2 Fig B), before parasite proliferation and translation (S3
187 Fig) were affected. High-throughput cDNA sequencing (RNA-Seq) revealed nearly
188 200 mRNAs that changed significantly in total and/or polysomal RNA fractions (S3
189 Table, sheets 3-7). 36% of the mRNAs that increased after RNAi were preferentially
190 expressed in procyclic forms, while 44% of those that decreased were preferentially
191 expressed in bloodstream forms (Table 1 and Fig 3 A). Increases in polysomal RNA
192 after RNAi were very slightly stronger than for total RNA. *RBP10* was the only mRNA
193 with strongly decreased polysomal loading after RNAi, whereas about 60 mRNAs
194 moved into polysomes (S3 Table sheet 3). 19 of these 60 mRNAs have a higher
195 ribosome density in procyclic forms than in bloodstream forms [11], suggesting that
196 their translation is inhibited in bloodstream forms. An example is the *COXVI* mRNA
197 (S3 Table sheet 3, S4 Fig A,B), which is known to be developmentally regulated
198 mainly at the level of translation [43]. Only 4 mRNAs had lower polysome occupancy
199 after RNAi (S1 Table sheet 3). Our results almost certainly under-estimate the extent
200 of translation changes, since we measured only whether an mRNA was in the
201 polysomal fraction or not. For example, if an mRNA had a few ribosomes before
202 RNAi, and more ribosomes after RNAi, this would not have been detected.

203 **Fig 3**

204 mRNAs bound to and/or regulated by RBP10.

205 Results for mRNA levels are displayed as proportional Venn diagrams made using
206 http://jura.wi.mit.edu/bioc/tools/venn.php and
207 http://www.eulerdiagrams.org/eulerAPE/

208 BS = Bloodstream form, PC = procyclic form. Categories were: Bound:
209 Supplementary Table S3 sheet 4. PC>BS and BS>PC: at least 2x significantly
210 regulated [10]. BS rbp10 down or up: affected by RNAi, Supplementary Table S1
211 sheets 6 & 7; PC+RBP10 down or up: affected by RBP10 expression in procyclics,
212 Supplementary Table S2 sheets 6 & 7. Fischer exact test values, calculated for the
213 dataset in the centre, are in blue italics.

214 A) Relationship between developmental regulation and the effects of RBP10
215 depletion in bloodstream forms

B) Relationship between developmental regulation and the effects of RBP10 expression in procyclic forms

C) Expression of RBP10 in procyclic forms related to the effects *rbp10* RNAi in bloodstream forms.

D) Developmental regulation of the mRNAs bound by RBP10.

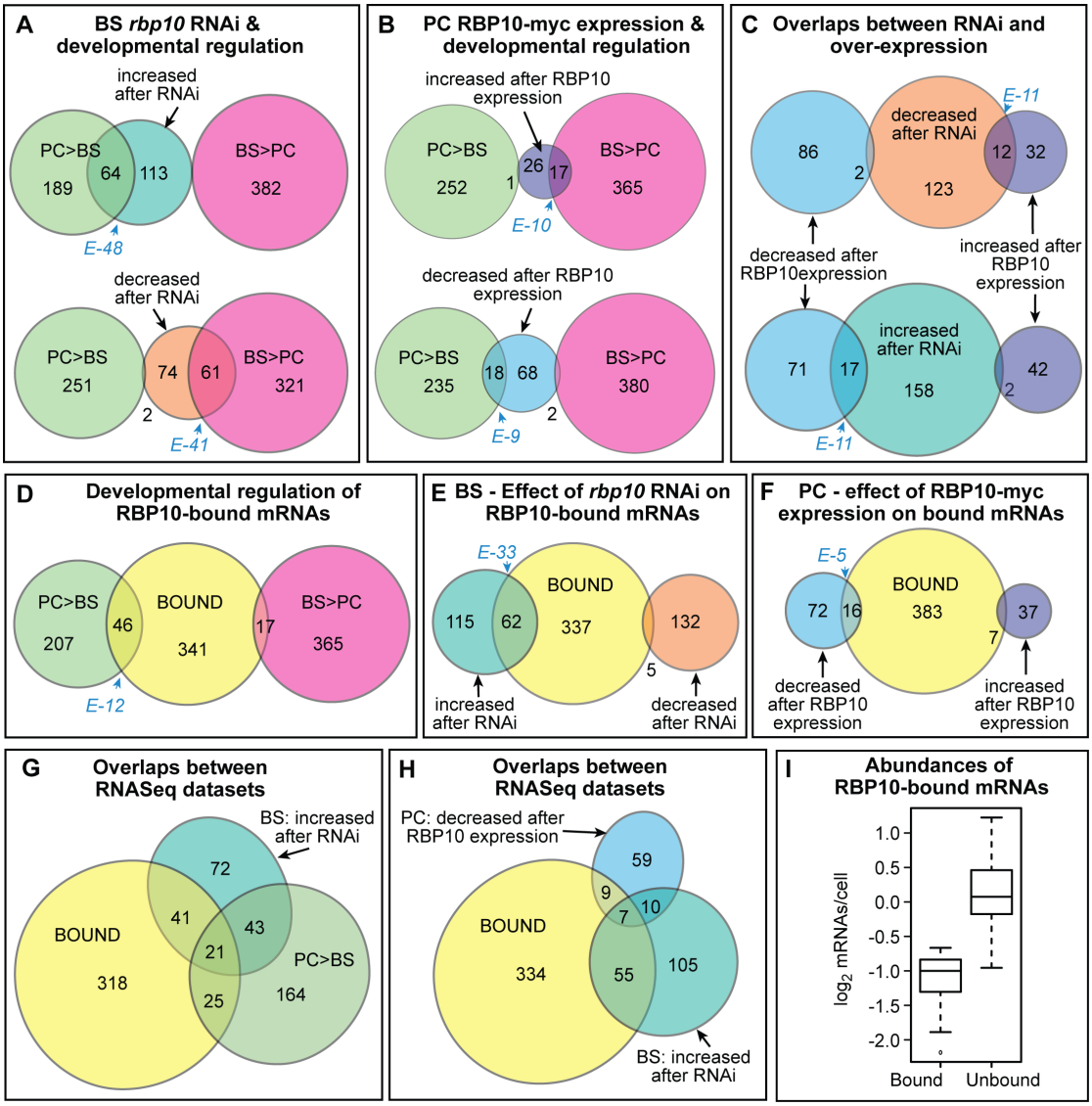
E) Effect of bloodstream-form *rbp10* RNAi on the mRNAs bound by RBP10.

D) Effect of procyclic-form RBP10 expression on the mRNAs bound by RBP10.

G) Overlaps between bound mRNAs, procyclic-specific mRNAs, and the mRNAs that increased after *rbp10* RNAi in bloodstream forms

H) Overlaps between bound mRNAs, the mRNAs that increased after *rbp10* RNAi in bloodstream forms, and the mRNAs that decreased after RBP10 expression in procyclic forms

I) Abundances of RBP10-bound mRNAs compared to all other mRNAs.



229

230

231 Next we analysed the effect of expressing C-terminally myc-tagged RBP10 (RBP10-
232 myc) in procyclic forms. After only 6h induction, (S2 Fig C, S4 Fig) we saw decreases
233 in procyclic-form-specific mRNAs, and increases in bloodstream-form specific
234 mRNAs (Fig 3 B and S4 Table). A subset of mRNAs showed opposite reactions in
235 the bloodstream RNAi and procyclic over-expression experiments, consistent with
236 being immediate RBP10 targets (Fig 3 C).

237 We had previously attempted to identify RBP10-bound mRNAs, but the methods that
238 we used had insufficient sensitivity [37]. Since a subsequent study showed that
239 RBP10 is indeed bound to mRNA [40], we made another attempt. This time, we used
240 our bloodstream-form cell line in which all RBP10 bears a cleavable tag (S2 Fig A),
241 and compared bound and unbound mRNAs by RNASeq. 400 mRNAs were identified
242 (S5 Fig A and S5 Table, sheet 3). 12% of them are more abundant in procyclic forms

than in bloodstream forms (Fig 3 D). Consistent with the tethering results, 15% of the RBP10 targets increased after *rbp10* RNAi in bloodstream forms, but only 1% decreased (Fig 3 E). Only 16 target mRNAs decreased after RBP10 expression in procyclic forms (Fig 3 F); this low number might be due to the very short duration of RBP10-myc expression. The RBP10-bound mRNAs that are more abundant in procyclic forms and increased after *rbp10* RNAi (Fig 3 G) are listed in Table 1: six of them are also decreased after RBP10 expression in procyclics (Fig 3 H). Scatter plots showing the effects of RNAi on bound RNA abundance and polysomal loading are shown in S5 Fig C and D.

Table 1

21 mRNAs that are bound to RBP10, are more abundant (RNA PC>BS) or better translated in procyclic forms than in bloodstream forms according to ribosome profiling (Riboprof PC>BS), and are increased after *rbp10* RNAi in bloodstream forms. For details see S5 Table sheet 8. The latter includes many additional mRNAs that fulfilled fewer conditions.

id	Annotation	**RNA PC>BS**	RiboProf PC>BS	de-creased by RBP10 in PC
Metabolism				
9.5900	glutamate dehydrogenase (GDH)	yes	yes	no
10.2350	pyruvate dehydrogenase complex	yes	yes	no
10.4280	Complex III subunit	yes	yes	no
1.4100	cytochrome oxidase subunit IV (COXIV)	yes	yes	no
11.15550	NADH-cytochrome b5 reductase,	yes	yes	no
4.4990	ubiquinol-cytochrome c reductase	yes	yes	no
7.210	proline dehydrogenase	yes	yes	no
9.4310	tricarboxylate carrier	yes	yes	no
9.7470	purine nucleoside transporter NT10	yes	yes	yes
11.6280	pyruvate phosphate dikinase (PPDK)	yes	yes	no
11.8990	cation transporter, putative	yes	yes	no
Surface				
10.10250	Procyclins #	yes	yes	yes
7.6850	trans-sialidase (TbTS)	yes	yes	yes
8.7340	trans-sialidase	no	yes	yes
4.3500	Amastin-like protein	yes	yes	no
RNA binding				
7.2660	ZC3H20	yes	no	no
7.2680	ZC3H22	yes	yes	yes
Unknown function				
9.13200	hypothetical protein	yes	yes	yes
11.16300	hypothetical protein	yes	yes	no
7.5550	hypothetical protein	yes	yes	no
10.3300	DUF4206 (cysteine-rich)	yes	yes	no
10.2360	hypothetical protein	yes	no	no

Overall, the RBP10-bound mRNAs are significantly less abundant than unbound mRNAs (Fig 3 I). At the proteome level [16], the products of RBP10-bound mRNAs are enriched in proteins that increase within 2-24h of the initiation of stumpy form differentiation, or early during conversion to procyclic forms. There is no enrichment for proteins that appear later in differentiation, or only in mature procyclics, and unregulated proteins or those that are decreased in procyclic forms are under-represented (S5 Table, sheet 6).

RBP10-bound mRNAs share a UAUUUUUU motif

A comparison of the 3'-untranslated region (3'-UTR) sequences of the most reliably RBP10-bound mRNAs with those of unbound mRNAs (S5 Table, sheet 5 and S1-S3 texts) revealed that the motif $UA(U)_6$ is highly enriched in the 3'-UTRs of RBP10 targets (Fig 4 A). It is present 1-9 times in 225 out of the 255 most strongly bound 3'-UTRs (Fig 4 B, S1 text). Transcriptome-wide, the number of $UA(U)_6$ motifs was strongly associated with RBP10 binding (Fig 4 C). Nevertheless, many mRNAs that contain $UA(U)_6$ were not pulled down with RBP10. Although there are some exceptions [44,45], RRM domains usually bind to single-stranded RNA [46-48]; the unbound motifs may have inappropriate conformations, or they might instead be bound by competing proteins.

Fig 4

RBP10 targets a UAUUUUUU motif

A) The 3'-UTRs of 188 mRNAs that were at least 3x enriched in both RBP10-bound samples (Supplementary Table S5 sheet 5) were compared with those of mRNAs that did not bind (<0.7x enrichment in both experiments) using DREME. Only 3'-UTRs annotated in tritrypDB were used (S1 text, S2 text, S3 text). The best-scoring motif found is shown.

B) Numbers of $UA(U)_6$ motifs in 255 manually annotated bound 3'-UTRs (S1 text, S5 Table sheet 5). One mRNA each had 7, 8 and 9 motifs. The numbers of motifs shown here are sometimes higher than with automatically annotated 3'-UTRs from the genome database, because some of the automatic 3'-UTRs are truncated.

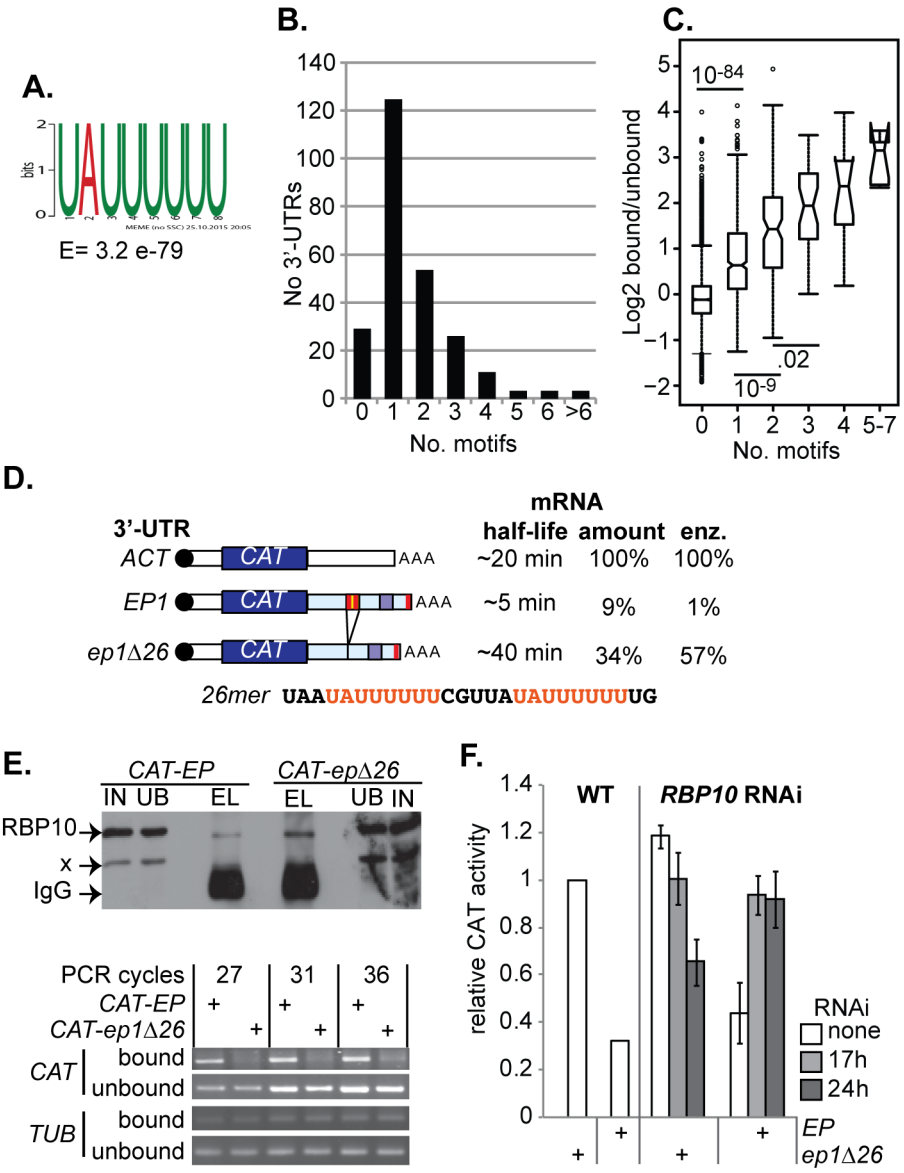
C) Effect of the $UA(U)_6$ motif on RBP10 binding: median ± 25th percentiles, with 95% confidence limits and outliers. Database 3'-UTRs were analyzed. Since these are often truncated, and are missing for 30% of genes, the numbers of motifs are under-estimated. Significant differences (Student t-test) are shown, For the mRNAs that

were enriched according to DeSeq (S5 Table, sheet 3), 284/353 annotated UTRs contained UA$(U)_6$.

D) Reporters used in (E) and (F) [34]. The red bars are UAUUUUUU elements, which are highlighted in the 26mer sequence.

E) RBP10 binding to the *EP* 3'-UTR requires the 26mer. Cells were UV-irradiated, and TAP-RBP10 was pulled down with IgG beads then released with TEV protease (upper panel, Western blot). RNAs were detected by reverse transcription and PCR using gene -specific primers (lower panel).

F) *rbp10* RNAi specifically affects expression of a reporter containing the 26mer. CAT activity was measured 24h after RNAi induction.



307 UA(U)$_6$ was previously identified in several mRNAs encoding procyclic-specific
308 proteins [43]. The 26mer 3'-UTR element that is required for *EP1* procyclin mRNA
309 instability [33,34] contains two UA(U)$_6$ motifs (Fig 4 B) and these are single-stranded
310 *in vivo* [49], as is required for RRM domain binding. Deletion of the region containing
311 UA(U)$_6$ from the COXV 3'-UTR increased protein expression without much effect on
312 the mRNA level [43], consistent with translation control. The mRNAs encoding
313 phosphoglycerate kinase (PGK) constitute a third example. In bloodstream-form
314 trypanosomes, most PGK is encoded by the gene *PGKC* and is targeted to
315 microbodies called glycosomes (S6 Fig A); in procyclic forms, a different gene,
316 *PGKB*, is expressed and the enzyme is in the cytosol [50,51] (S6 Fig C). The *PGKB*
317 *and PGKC* coding regions are identical except for a C-terminal targeting sequence in
318 *PGKC*, so they were not distinguished in our initial RNASeq analysis. However, the
319 3'-UTRs are completely different, and determine developmental regulation [52]. The
320 UA(U)$_6$ motif is present twice within a 3-UTR segment that is required for suppression
321 of PGKB expression in bloodstream forms [53] (S1 text). Using our RNASeq pull-
322 down data, we compared the abundances of several 15nt sequences specific to the
323 *PGKB* and *PGKC* mRNAs and found that that *PGKB* mRNA is almost exclusively in
324 the bound fraction, whereas most *PGKC* mRNA is unbound. A similar analysis gave
325 RBP10 binding ratios of 2.8:1 for *EP1* and 1.3:1 for *GPEET* mRNAs. The *GPEET*
326 mRNA 3'-UTR has UA(U)$_5$ repeats, but no UA(U)$_6$, and it is not yet clear whether it is
327 bound by RBP10.

328 To test whether the 26mer was required for binding of the *EP1* 3'-UTR to RBP10, we
329 used cell lines that expressed chloramphenicol acetyltransferase (CAT) reporter
330 mRNAs with either the intact wild-type *EP1* 3'-UTR, or a version without the 26mer
331 (*ep1*Δ*26*) (Fig 4 D) [34]. Co-immunoprecipitation of the *CAT* mRNA with RBP10
332 depended on the presence of the 26mer (Fig 4 E). Depletion of RBP10 by RNAi also
333 increased expression from the *CAT-EP* reporter but not from *CAT-ep1*Δ*26* (Fig 4 F).
334 These results demonstrated both that RBP10 binds to the *EP* 3'-UTR via the 26mer
335 sequence, and that the 26mer is necessary for RBP10-mediated regulation of an
336 mRNA containing the *EP1* 3'-UTR.

337 **Expression of RBP10 determines bloodstream form identity**

338 To test the role of RBP10 in differentiation, we first examined conversion from the
339 bloodstream form to the procyclic form. Our monomorphic bloodstream forms with
340 *rbp10* RNAi, or treated with the differentiation-inducer *cis* aconitate, are unable to
341 multiply under procyclic culture conditions. We therefore used differentiation-

342 competent trypanosomes (EATRO 1125, Antat1.1 strain). As expected, growth of
343 these cells at high density in methyl cellulose [3] for three days resulted in expression
344 of the a -form marker PAD1 [54]. RBP10 was also reduced (Fig 5 A, lanes 2 &3),
345 consistent with its absence in the stumpy-form proteome [54]. The cells were now
346 treated with 6mM cis-aconitate in bloodstream-form medium at 27°C for 17h. Upon
347 transfer to procyclic medium without cis-aconitate, the cells started to grow as
348 procyclics within 24h (Fig 5B).

349

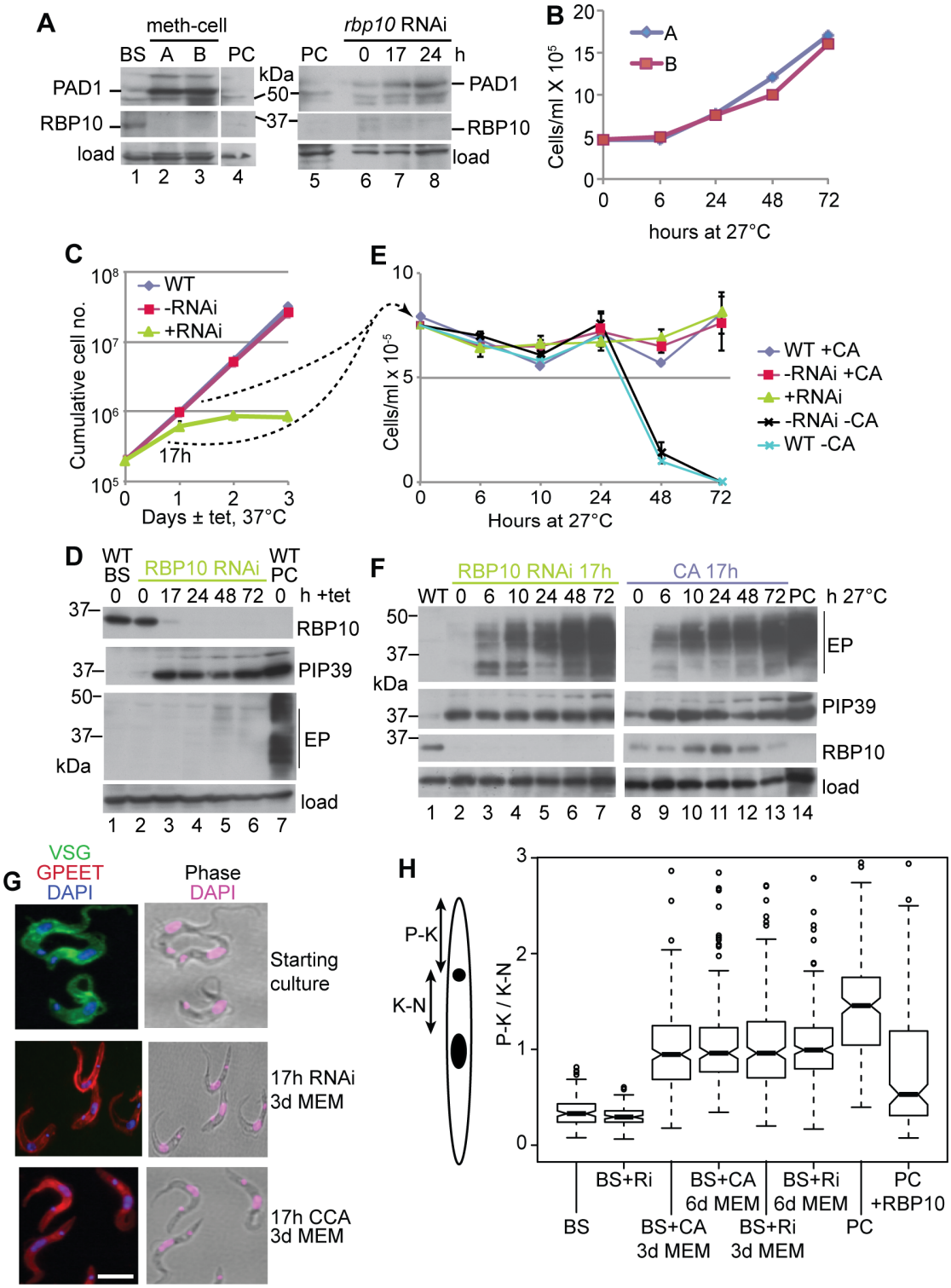


Fig 5

RBP10 depletion primes bloodstream forms for differentiation to procyclic forms

A) Expression of PAD1 and RBP10 in trypanosomes incubated at maximal density in methyl cellulose-containing medium for 3 days (left panel), or after *rbp10* RNAi (right

355 panel). The RBP10 signal is weak, probably because the samples were not
356 denatured in order to allow PAD1 detection.

357 B) Growth of trypanosomes incubated at maximal density in methyl cellulose-
358 containing medium for 3 days, followed by cis-aconitate treatment for 17h at 27°C.
359 Time 0 is the time of transfer to procyclic conditions.

360 C) Cumulative growth curve of bloodstream-form (BS) trypanosomes with and
361 without RNAi. WT = wild-type (*tet* repressor only); -RNAi: RNAi cell line with no
362 tetracyline: +RNAi: RNAi cell line with tetracycline.

363 D) Expression of RBP10, PIP39, and EP procyclin were measured by Western
364 blotting up to 3d after tetracycline induction of RNAi.

365 E) 17h after RNAi induction at 37°C in bloodstream form medium (+RNAi), or
366 incubation at high density in procylic medium with 6mM cis-aconitate at 27°C (+CA),
367 trypanosomes were placed in procyclic medium (MEM-pros) at 27°C with neither
368 tetracycline nor cis-aconitate. The cell densities of the cultures were monitored.
369 Negative controls were either wild-type or uninduced RNAi lines, without cis-
370 aconitate. Cell densities are shown; they started to increase at 72h.

371 F) Protein expression after transfer to procyclic medium at 27°C after 17h pre-
372 treatment as in (E).

373 G) Immunofluorescence: Cell morphology and expression of GPEET procyclin after
374 3d cultivation in procyclic medium. DNA is stained with DAPI.

375 H) The cartoon shows the way in which the position of the kinetoplast, relative to the
376 nucleus (P-K/K-N), was measured. The box plot shows the median with 25th
377 percentiles and 95% confidence limits. Some outliers have been deleted. The full
378 results, together with the numbers of trypanosomes measured, are in S7 Fig A.

379 BS - normal bloodstream forms; PC- normal procyclic forms; CA - 17h cis-aconitate
380 pretreatment; RNAi or Ri; 17h *rbp10* induction; +RBP10 - induced expression or
381 RBP10 for 2 days. There were no significant differences between the 3-day and 6-
382 day differentiating populations. All other pairs were significantly different (Student t-
383 test, $p<0.05$).

384 Induction of *rbp10* RNAi in differentiation-competent cells in log phase resulted in
385 growth inhibition (Fig 5 C), weak PAD1 expression (Fig 5 A, lanes 6-8) and, within
386 17h, expression of the differentiation-regulating phosphatase PIP39 [55] (Fig 5 D).
387 After 17h RNAi, transfer of the cells to procyclic-form medium at 27°C (without cis-
388 aconitate) allowed them to survive and eventually start to grow (Fig 5 E), with

389 expression of EP procyclin (Fig 5 F, lanes 1-7) and GPEET procyclin (Fig 5 G)
390 proteins. As expected, cells without RNAi induction died (Fig 5 E). The RNAi
391 experiment had to be done without high density pre-incubation in methyl cellulose
392 since RNAi can only be induced in growing trypanosomes. As a control, we therefore
393 grew differentiation-competent cells to maximum density in liquid medium, and
394 treated them with cis-aconitate for 17h at 27°C [56] before transfer to procyclic
395 medium without cis-aconitate, also at 27°C. In these cells, weak RBP10 expression
396 persisted for 2 days (Fig 5 F, lanes 8-13) but the growth and procyclin expression
397 kinetics were similar to those seen after *rbp10* RNAi (Fig 5 E, F, G).

398 The rather slow resumption of growth after both *rbp10* RNAi and 17h cis-aconitate
399 treatment, and the persistence of RBP10 in the cis-aconitate experiment, suggested
400 that the population might be a mixture of differentiation-competent and incompetent
401 cells. Indeed, both differentiating cultures were mixtures of normal-looking,
402 proliferating, and dying or abnormal forms, and even after 3 days in procyclic culture,
403 about 9% of the cells in both cultures retained bloodstream-form morphology and
404 PGK localization (S6 Fig D). In bloodstream forms, the kinetoplast is at the posterior
405 end of the cell (Fig 5 G), which means that the distance between the posterior and
406 the kinetoplast (P-K) is much less than the distance from the kinetoplast to the
407 nucleus (K-N) (Fig 5 H and S7 Fig A, B; "BS"). Procyclic forms are longer and the
408 kinetoplast is nearer to the nucleus (Fig 5 G), giving a much higher P-K/K-N ratio (Fig
409 5 H and S7 Fig A, B; "PC"). The differentiating cultures showed very heterogeneous
410 P-K/K-N ratios, with medians that were intermediate between those of bloodstream
411 and procyclic forms (Fig 5 H, S7 Fig A-C). As expected, cells with low ratios typical of
412 bloodstream forms were preferentially GPEET negative (S7 Fig D,E).

413 We next inducibly expressed RBP10-myc in differentiation-competent procyclic
414 forms. For the results shown, we made bloodstream forms with inducible RBP10-
415 myc, differentiated them into procyclic forms, grew these for more than 3 months,
416 and then induced RBP10 expression. However similar results were obtained by
417 differentiating bloodstream forms into procyclic forms, and then making three new
418 cell lines with the inducible RBP10-myc plasmid. After RNAi induction, cell growth
419 was inhibited, as expected [37] (Fig 6 A). After 48h, expression of procyclin was
420 strongly reduced, alternative oxidase increased, and untagged RBP10 was detected
421 (Fig 6 B, C). At this point the cell population was very heterogeneous. Some cells
422 had normal procyclic morphology and expressed EP and GPEET procyclin; some
423 were clearly abnormal; while another subset had clear bloodstream-form morphology
424 (Fig 6D). This was reflected in the P-K/K-N ratios, which ranged from procyclic to

bloodstream-form patterns (Fig 5 G, S7 Fig F). Flow cytometry analysis revealed that about 50% of cells had decreased, or absent, procyclin expression (Fig 6 E, F). Cells with bloodstream-form morphology were usually (but not always) procyclin negative (S7 Fig G). Notably, *VSG* mRNA was present (Fig 6 C). We do not know which VSGs are expressed in the bloodstream-form-morphology cells, but a subset of them stained with a polyclonal antibody against VSG117 (Fig 6 D, G, H). PGK staining in the bloodstream-form-like cells seemed reduced and somewhat more punctate than in procyclic forms, but had not yet assumed a purely glycosomal pattern (S6 Fig E). Remarkably, the RBP10-induced cells could now survive only if they were transferred to bloodstream-form growth conditions (Figure 6A). After 6 days in these conditions, all surviving cells had bloodstream-form morphology and 60% of them cross-reacted with the anti-VSG117 antibody.

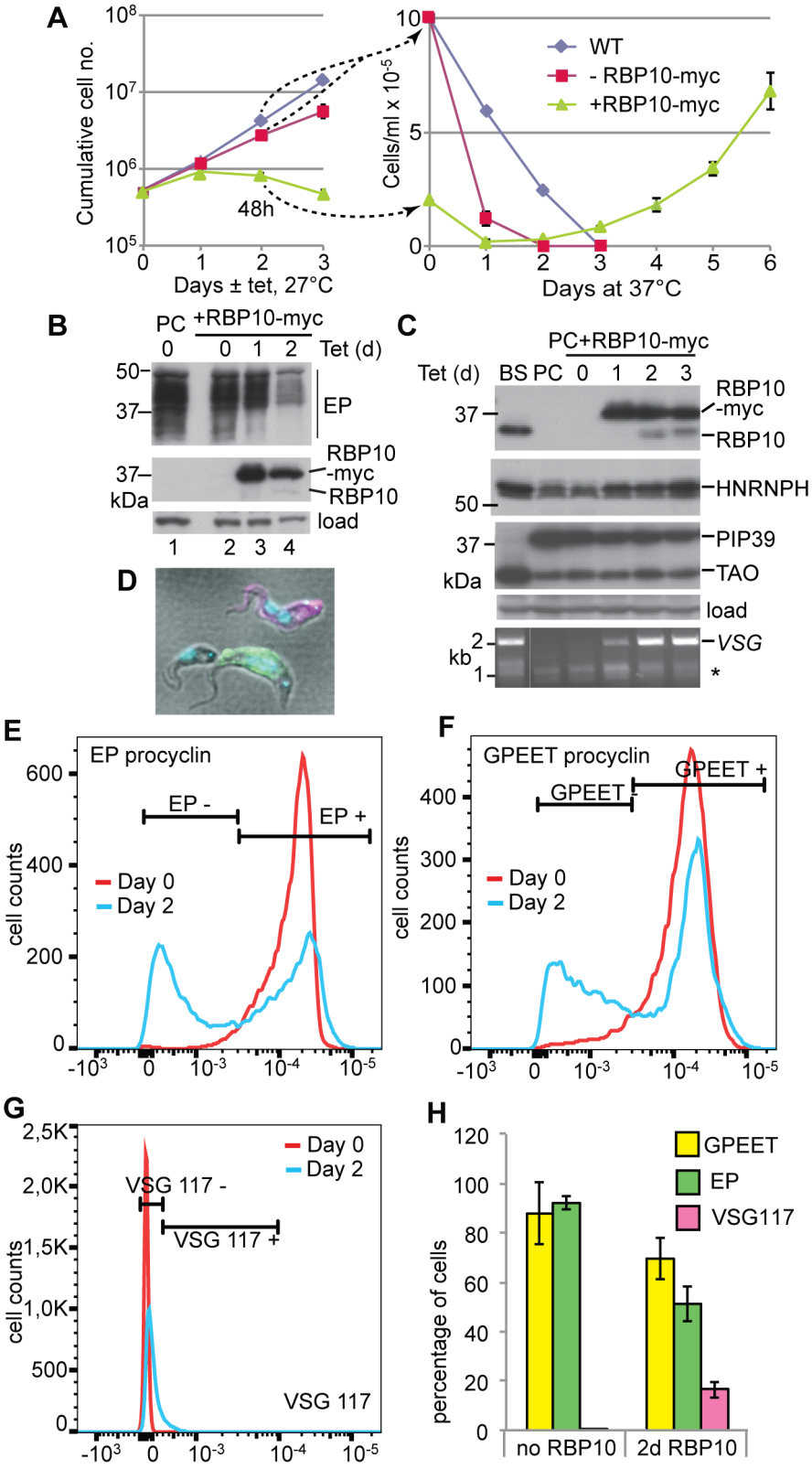


Fig 6

RBP10 expression in procyclic forms causes differentiation to bloodstream forms

A) Cell counts in a typical experiment. Expression of RBP10-myc was induced using tetracycline at 27°C in procyclic-form medium. The left-hand panel is a cumulative growth curve. After 48h cells were transferred to bloodstream-form medium at 37°C; in the right-hand panel cell densities are shown.

B) Expression of EP procyclin and RBP10 after induction of RBP10-myc expression. Cells were cultured in procyclic medium at 27°C. Expression of EP Procyclin protein decreased after 48h of RBP10-myc induction, while untagged RBP10 becomes detectable.

C) As in (B) except that culture continued for 4 days. Proteins that are expressed more in bloodstream forms than in procyclic forms were detected by Western blotting. TAO is trypanosome alternative oxidase; HNRNPH is an RNA-binding protein. *VSG* mRNA was detected by RT-PCR using a spliced leader primer and a primer that hybridizes to a conserved region within the *VSG* 3'-UTR.

D) Trypanosomes were taken after 48h RBP10-myc induction. They were stained for EP procyclin and phosphorylated GPEET procyclin (green), VSG117 (magenta) and DNA (cyan). The stain is overlaid with differential interference contrast (grey). The panel shows a typical green procyclic form, a bloodstream form with two kinetoplasts, and a third cell which resembles a bloodstream form (terminal kinetoplast) but was not stained by anti-VSG117. This cell may express a VSG that does not react with the anti-VSG117 antibodies.

E) Quantitation of surface EP procyclin by flow cytometry, comparing cells with (day 2) and without (day 0) induced RBP10-myc expression.

F) Quantitation of surface phospho-GPEET procyclin by flow cytometry, comparing cells with (day 2) and without (day 0) induced RBP10-myc expression.

G) Quantitation of staining with anti-VSG117 by flow cytometry, comparing cells with (day 2) and without (day 0) induced RBP10-myc expression.

H) Measurement of surface protein expression by FACSscan. Values for the different windows are mean and standard deviation from 4 replicates. Similar results were obtained by manual counting of stained smears.

The conversion of procyclic forms to bloodstream forms was highly reproducible. It was not due to persistence of a few bloodstream forms in the procyclic cultures, since survival always depended on induced RBP10 expression (Fig 6 A). RBP10 induction for two days was required; this was also the time needed to detect native RBP10 expression by Western blotting. Moreover, our preliminary results indicate that transient expression of RBP10 from an episome is sufficient to generate bloodstream forms. In contrast, attempts to obtain bloodstream forms by RBP10 expression in Lister 427 cells that have been in procyclic form culture for more than two decades were not successful.

Interestingly, epimastigote-like cells were never seen after induction of RBP10-myc expression in our differentiation-competent cells (24h or 48h), and the epimastigote-specific surface protein BARP was not detected by Western blotting or immunofluorescence. Although we may have missed transient formation of epimastigotes, this is unlikely since the behaviour of the trypanosomes was clearly not synchronous. Moreover, the presence of dividing cells with bloodstream-form morphology, suggested that RBP10 induction might be causing a direct conversion of procyclic forms to bloodstream forms, "jumping" past the non-dividing metacyclic form. We therefore compared the mRNA changes after 6h RBP10-myc induction with those found in salivary gland parasites (a mixture of epimastigotes and metacyclics) (S4 Table, sheets 5 and 6). The transcriptome comparisons and morphological analysis suggested that after 2 days of RBP10 expression the cell population might include both metacyclic and bloodstream forms. Interestingly, however, the mRNA encoding the meiosis protein MND1, which is specific to a pre-metacyclic intermediate form, is bound by RBP10 and *decreases* upon RBP10 induction. Evidence so far thus suggests that RBP10 induction causes a developmental switch that by-passes the epimastigote stage.

Discussion

Our results indicate that the presence of RBP10 defines the identity of a trypanosome as a replication-competent bloodstream form.

RBP10 binds to procyclic-specific mRNAs with a $UA(U)_6$ motif in the 3'-UTR, suppressing translation and causing mRNA destruction. However, many RNAs that are bound by RBP10 did not change in either abundance, or in the percentage in polysomes, after RBP10 RNAi or forced expression. As noted above, our assay did not measure ribosome density, so it is likely that many changes in translation went

504 undetected. However, our results are completely consistent with work on other RNA-
505 binding proteins: usually, only some of the bound mRNAs change after the protein is
506 removed (e.g. [57]). For most mRNAs, many different proteins can bind along the
507 same 3'-UTR - a sequence of 300nt is likely to bind at least twenty proteins - and it is
508 their combinatorial effects that determine mRNA behaviour.

509 RBP10 target mRNAs include several that have been studied in considerable detail,
510 including those encoding cytochrome complexes and the procyclins. However, the
511 direct effects of RBP10 are by no means sufficient to explain its far-reaching
512 influence, since many mRNAs that change in translation or abundance after RBP10
513 depletion are not bound by RBP10. This suggests that RBP10 depletion may initiate
514 a cascade of events. Our results show how this could occur. Notably, the direct
515 RBP10 targets include the mRNAs encoding three potential RNA-binding proteins,
516 ZC3H20, ZC3H21, and ZC3H22, These proteins are normally expressed only in
517 procyclic forms [15,16,58]. ZC3H20 binds to, and stabilises, at least two procyclic-
518 specific mRNAs [58], of which one, *MCP12*, is not bound by RBP10 but was
519 increased after *rbp10* RNAi. ZC3H22 suppresses expression in the tethering assay
520 [39], and is required for proliferation of procyclic forms [15,42]. Suppression of
521 ZC3H21 and ZC3H22 expression is clearly very important to bloodstream-form
522 trypanosome survival: their mRNA 3-UTRs contain, respectively, no fewer than 5 and
523 7 RBP10 binding motifs.

524 In addition to controlling ZC3H20-22, RBP10 binds to and represses mRNAs
525 encoding two procyclic-specific protein kinases (Tb927.8.6490, Tb927.11.15010) and
526 one procyclic specific protein phosphatase (Tb927.10.8050). Also bound by RBP10,
527 but with no detected effect of RNAi, are the transcripts encoding and protein kinases
528 RDK1 and NRKA. NRKA [15] is up-regulated after the onset of stumpy form
529 differentiation; it is possible that RBP10 indeed regulates its translation but we did not
530 detect this because of the insensitivity of our assay. RDK1 is a differentiation
531 repressor [59] and is more abundant in bloodstream forms, making regulation by
532 RBP10 unlikely. Similar to RBP10, depletion of *RDK1* primes bloodstream forms to
533 differentiate to procyclic forms. More than 20 mRNAs show similar changes after
534 RNAi targeting either *RDK1* or *RBP10*, and 16 of these are both bound and regulated
535 by RBP10. Surprisingly, the *RBP10* transcript was not significantly affected after
536 *RDK1* RNAi; however, the protein level and phosphorylation status of RBP10 were
537 not determined.

538 The changes in RNA regulators, protein kinases and phosphatases, and perhaps
539 also the changes in metabolism that are directly induced by RBP10 loss (e.g. [60])

will trigger downstream effects on signaling pathways. Such secondary effects of *rbp10* RNAi must include the increase in PIP39 (Tb927.9.6090), a key regulator of differentiation [55], and decreases in mRNAs encoding the RNA-binding proteins PUF11, RBP9, ZC3H31, ZC3H46, ZC3H48, DRBD5 and HNRNPH (S1 Table, sheet 7). These in turn will trigger further changes in gene expression.

So far, the only condition known to allow differentiation of *T. brucei* procyclic forms to bloodstream forms *in vitro* is induced expression of the putative RNA-binding protein RBP6. This mimics the natural differentiation pathway, with formation of epimastigotes within 24h and appearance of VSG-expressing metacyclic forms 4-5 days later [61]. *RBP6* mRNA has low abundance in both procyclic an bloodstream forms, and is maximal in salivary gland parasites [38]. RBP6 action is therefore transient, although it might possibly define the epimastigote form; its mechanism of action is unknown. Depletion of the RNA-binding protein DRBD18 from procyclic forms caused a partial switch to a bloodstream-form or epimastigote expression pattern; the mRNAs that increased included those encoding RBP6 and RBP10, but the ability to differentiate further was not tested [62]. Heat shock can also cause small increases in epimastigote-specific transcripts [63]. The rapid conversion of procyclic forms to bloodstream forms by RBP10 expression, in contrast, might be regarded as a form of *trans*-differentiation since the intermediate epimastigote stage was not detected. In animal cells, *trans*-differentiation can be achieved through changes in expression of transcription factors, either by direct manipulation or by changing expression of microRNAs that control transcription factor expression [64-66]. However, these changes have so far always been unidirectional. *Trans* differentiation that works in both directions according to the presence or absence of a single protein has, to our knowledge, not previously been observed.

It has long been known that stumpy forms cannot revert to long-slender bloodstream forms, and recent results indicate that the differentiation of stumpy forms to procyclics is an irreversible bi-stable switch [15]. We suggest that expression of RBP10 governs a bi-stable switch during the transition from bloodstream forms to stumpy forms, and from procyclic forms to bloodstream forms. The bi-stable character can be explained by the fact that RBP10 directly suppresses procyclic-specific post-transcriptional regulators. While RBP10 is present, the bloodstream-form state is self-reinforcing; loss of RBP10 enables procyclic regulators to be expressed, while expression of RBP10 in procyclic forms tips the balance towards a bloodstream-form expression pattern. In conclusion, our results demonstrate how differentiation states can be maintained by a post-transcriptional cascade.

Materials and methods

Parasite culture and plasmid constructs

T. brucei Lister 427 [67] and pleomorphic AnTat 1.1 (EATRO 1125) [56] cells expressing the Tet-repressor were used. Lister 427 were used for all experiments except those shown in Figs 5 and 6. Bloodstream form parasites were cultured in HMIi-9 medium supplemented with 10% fetal bovine serum at 37^{o}C with 5% CO2. The procyclic forms were grown at 27^{o}C in MEM-Pros medium supplemented with 10% heat inactivated fetal bovine serum. Stable cell lines were generated and maintained as described in [56,67] except that selection of the pleomorphic cells was done with 8µg/ml hygromycin and 2µg/ml blasticidin.

The tetracycline inducible constructs for *RBP10* RNAi and over expression were described in [37]. All new constructs and oligonucleotides are in S6 Table. A bloodstream-form cell line expressing all RBP10 with a tandem affinity purification tag at the N-terminus was generated by inserting the tag at the 5' end of one endogenous *RBP10* open reading frame (pHD2506) and deletion of the second copy (pHD2061). Plasmids for the tethering assays are in S6 Table. The tethering constructs were separately transfected in a cell line constitutively expressing the *CAT* reporter with 5 copies of *boxB* preceding the *ACT* 3' UTR. Expression of the fusion protein was induced for 24h using tetracycline (100ng/ml) and the CAT assays carried out as described in [25].

Cross-linking and RNA immunoprecipitation

3x10^{9} bloodstream form cells expressing TAP-RBP10 were irradiated using UV (400 mj/cm2), washed in cold PBS and the cell pellet snap frozen in liquid nitrogen. The RNA immunoprecipitation was done as described in [24]. Briefly, the extracts were incubated with the beads, and the unbound fraction was collected. After washing, bound RBP10 was eluted using TEV protease. After digestion with 50µg proteinase K at 42°C for 15 minutes to reduce cross-linked protein, RNA was isolated from both the bound and unbound fractions using Trifast reagent (Peqlab, GMBH). To assess the quality of the purified RNA, an aliquot of the sample was analysed by Northern blotting and the blot hybridized with a splice leader probe. Total RNA from the unbound fraction was depleted of ribosomal RNA (rRNA) using RNAse H as described in [68] except that a cocktail of 50-base DNA oligos complementary to trypanosome rRNAs was used [63]. The recovered RNA from both bound and unbound samples was then analysed by RNA-Seq.

Tandem affinity purification (TAP) and genome-wide yeast two-hybrid screen

Approximately $2x10^{10}$ bloodstream form cells expressing in-situ TAP-RBP10 were harvested and used for TAP as previously described [69]. Three biological replicates were done, with or without RNAse A treatment; triplicate results from a cell line with inducible GFP-TAP served as the background control. The eluate was separated on a 10% SDS-polyacrylamide gel for only 2 cm and the proteins visualized by colloidal Coomassie staining. The gel area was excised and analysed by mass spectroscopy. For the yeast two-hybrid screen, the RBP10 open reading frame was cloned into the pGBKT7 plasmid (pHD2361). The construct was used as a bait to screen a library made up of *T. brucei* random genomic DNA fragments. The screening, high throughput sequencing and analysis were done as previously described [70].

Polysome fractionation

Approximately $3\text{-}5x10^{8}$ cells were collected by centrifugation (850 x g, 10 minutes, 21°C), and then treated in serum free media with 100µg/ml cycloheximide for 7 minutes at room temperature. The pellet was washed in 1 ml ice cold PBS, lysed in 350µl lysis buffer (20mM Tris pH7.5, 20mM KCl, 2mM MgCl2, 1mM DTT, 1000u RNasin (Promega), 10µg/ml leupeptin, 100µg/ml cycloheximide, 1× complete protease Inhibitor without EDTA (Roche), 0.2% (vol/vol) IGEPAL) by passing 15-30 times through a 21G needle, followed by centrifugation (15000 x g, 10 minutes, 4°C) to clear the lysate. KCl was adjusted to 120 mM and the clarified lysate loaded on top of a 4 ml continuous linear 15-50% sucrose (w/v) gradient in polysome buffer (20mM Tris pH7.5, 120mM KCl, 2mM MgCl2, 1mM DTT, 10µg/ml leupeptin, 100µg/ml cycloheximide). After 2 hours of ultracentrifugation (40000 rpm, 4°C; Beckman SW60 rotor), 400µl fractions were collected using Teledyne Isco Foxy Jr. gradient fractionator system and RNA was isolated using Trifast reagent (Peqlab, GMBH). In the case of RBP10 tethering, lambdaN-RBP10 was induced (24h) and the distribution of the *CAT* reporter and *alpha-tubulin* mRNAs in the collected fractions detected by Northern blotting as described in [71].

For RNASeq, bloodstream form cells plus or minus *RBP10* RNAi for 15h and, the procyclic form cells with or without RBP10-myc over-expression for 6h were used. In this case 250µl fractions were collected from each gradient and RNA was isolated after pooling the fractions into two groups; i) lighter fractions including monosomes, ii) denser fractions with at least two ribosomes. The amount of mRNA in the pooled samples was assessed by Northern blotting using splice leader RNA as probe. Also, total RNA was prepared from the input samples (~10% of total cell lysate) to quantify

changes in the steady state mRNAs levels. All samples were depleted of rRNAs (as above) prior to analysis by RNA-seq.

High throughput RNA sequencing and bioinformatic analysis

RNA-seq was done at the CellNetworks Deep Sequencing Core Facility at the University of Heidelberg. For library preparation, NEBNext Ultra RNA Library Prep Kit for Illumina (New England BioLabs Inc.) was used. The libraries were multiplexed (6 samples per lane) and sequenced with a HiSeq 2000 system, generating 50 bp single-end sequencing reads.

The quality of the raw sequencing data was checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and the sequencing primers removed using Cutadapt [72]. The data was aligned to the *T. brucei* TREU 927 reference genome using Bowtie [73], then sorted and indexed using SAMtools [74]. Reads aligning to open reading frames of the TREU 927 genome were counted using custom python scripts. Analysis for differentially expressed genes was done in R using the DESeq package [75] with an adjusted p-value cut-off of 0.05. Comparative analysis was limited to a list of unique genes modified from [9]. Gene annotations are manually updated versions of those in TritrypDB (http://tritrypdb.org/tritrypdb/), and categories were assigned manually. Statistical analysis was done in R. The 3' UTR motif enrichment search was done using DREME [76]; annotated 3' UTRs sequences were downloaded from tritrypDB and we considered only the mRNAs with 3' UTR >20 nt. Manual 3'-UTR annotation was done using the RNASeq reads and poly(A) site data [9,11,77] in tritrypDB.

Analysis of 3'-UTR reporters

Bloodstream form cells expressing the *CAT* reporter with either full length *EP1* 3' UTR (pHD1610) or a mutant version (pHD1611) lacking the 26mer (*EP1*Δ*26*) instability element [34] were used for RNA immunoprecipitation . For the pull down, $4x10^8$ cells (without the stem-loop) were irradiated using UV (400 mj/cm^2), washed in cold PBS and the cell pellet lysed in 350µl lysis buffer (10mM Tris pH 7.5, 10mM NaCl, 1000u RNasin (Promega), 1× complete protease Inhibitor without EDTA (Roche), 0.1% IGEPAL) by passing 15 times through a 21G needle. The lysate was cleared by centrifugation at 15000g for 10 minutes at 4°C, NaCl was adjusted to 150mM followed by incubation with 50µl anti-RBP10 [37] coupled agarose beads for 2 hours at 4°C. After washing the beads 5 times with IPP150 buffer (10mM Tris pH 7.5, 150mM NaCl, 0.1% IGEPAL), the beads were treated with 20µg proteinase K at 42°C for 15 minutes and RNA was isolated from both bound and unbound fractions

using Trifast reagent (peqlab, GMBH). Equal amounts of the recovered RNA (eluate and flow through fractions) were reverse-transcribed using RevertAid First Strand cDNA Synthesis Kit (Thermal Scientific) according to the manufacturer's instructions. 2µl of the cDNA was used as template in a 50µl PCR reaction to detect the *CAT* and *alpha-tubulin* genes. PCR was done using Q5 DNA polymerase and buffer (NEB) with 0.5µm of the following primers, for *CAT* (CZ5725; CZ689) and *alpha-tubulin* (CZ5725; CZ6168); the forward primer is the same for both since it anneals to the splice leader. Aliquots (10µl) were removed after 27, 31 and 36 cycles and analysed by agarose gel electrophoresis.

To determine if the regulation of the *EP* mRNA by RBP10 depends on the 26mer instability element, a stem-loop construct (pHD1984) targeting *RBP10* was transfected into the two cell lines [37]. *RBP10* RNAi was induced using 100ng/ml tetracycline for 17h or 24h and CAT activity was measured.

RBP10 and trypanosome differentiation

For bloodstream-procyclic form conversion, pleomorphic AnTat 1.1 bloodstream form cells with stem-loop RNAI (pHD1984) targeting *RBP10* were used. 17h after RNAi induction, the cells were pelleted, resuspended (~$8x10^5$ cells/ml) in procyclic form (MEM-pros) medium and incubated at 27°C. As positive controls, wild type or uninduced cells ($2x10^6$ cells/ml) were treated with 6mM cis-aconitate (Sigma) at 27°C; after 17h the cells were transferred into procyclic form media (~$8x10^5$ cells/ml) and maintained at 27°C. Samples were taken at different times for Western blotting (~$5x10^6$) and for morphological analysis.

To convert procyclic to bloodstream forms, AniTat 1.1 bloodstream form cells with an inducible RBP10-myc construct (pHD2098) were differentiated to procyclic forms using cis-aconitate as described above. The cells were cultured in presence of hygromycin (8µg/ml) and phleomycin (0.2µg/ml) for more than 3 months to generate well-established procyclic forms. RBP10-myc was induced using 100ng/ml tetracycline. Marker proteins were detected by Western blot. VSG transcripts were detected by semi-quantitative RT-PCR as previously described [61] using primers CZ6308/CZ6309. To obtain growing bloodstream forms, RBP10-myc was induced for 48 hours, the cells were pelleted, resuspended ($2x10^5$ cells/ml) in HMI-9 medium and then incubated at 37°C with 5% CO_2. The cell density was monitored for 6 days or more; wild type or uninduced cells served as control.

Morphological analysis

To stain surface proteins, dried smears were fixed with 100% methanol at -20^{o}C for 15 minutes, rehydrated in 1x PBS for 15 minutes, blocked for 20 minutes with 20% FCS, then labeled with mouse anti-EP (1:500; Cedarlane, Canada), mouse anti-phospho-GPEET (1:500; Cedarlane, Canada) and rabbit anti-VSG-117 (1:500, from either G. Cross or P. Overath). Anti-BARP staining was done as described in [5], using a procyclic cell line with the *BARP* open reading frame in the procyclin locus as a positive control. For PGK staining, 2x10^{6} cells were collected by centrifugation, re-suspended, fixed in 4% paraformadehyde (in 1x PBS) for 18 minutes, sedimented again for 2 minutes, re-suspended in PBS and allowed to settle on poly-L-lysine coated slides for 30 minutes. Before staining, slides were blocked with 20% fetal calf serum (1x PBS) for 20 minutes. Cells were permeabilised with 0.2% (v/v) Triton-X 100 (in 1xPBS) for 15 minutes at room temperature, washed twice then incubated for one hour with rabbit anti PGK antibody (1:1500, [78]). The cells were washed 3 times before being stained with fluor-conjugated secondary antibody (1:500; mouse-Cy3 or rabbit-Alexa-488; Molecular probes, Eugene). Cellular DNA was stained with 100ng/ml DAPI in 1x PBS for 15 minutes. Images were taken using Olympus Cell-R microscope. Random fields were photographed (by E.M.) and analysed (by C.C.) using Fiji [79]. We attempted blind analysis but this was not possible because the differences between samples were too obvious.

Flow cytometry

Approximately 5x10^{6} cells were fixed with 2% formaldehyde/0.05% glutaraldehyde at 4^{o}C for at least 1 hour. The cells were pelleted, washed twice with PBS then incubated with 200µl (2% BSA in PBS) mouse anti-EP (Cedarlane, Canada; 1:500), anti phosph-GPEET (Cedarlane, Canada; 1:500), or rabbit anti-VSG-117 (1:500) for one hour on ice. After washing twice, the cells were stained with the secondary antibody (1:500; mouse-Cy5 or rabbit-Alexa-488; Molecular probes, Eugene) for one hour. Cells stained only with the secondary antibody and the unstained cells served as the negative controls. Flow cytometry was performed with BD FACSCanto II flow cytometer and the data analysed using FlowJo software (TreeStar Inc.).

Western blotting

Western blotting was done as previously described [71]. Antibodies were: anti-RBP10 (rat, 1:2000, [37]), anti-HNRNPH (rabbit,1:5000, [20]), anti-TAO (rabbit,1:100, [78]), anti-EP (mouse,1:2000, Cedarlane, Canada), anti-phospho-GPEET (mouse,1:2000, Cedarlane, Canada), anti-BARP (rabbit,1:2500, [5]), and anti-PIP39

(rabbit,1:1000, [55]) and anti-PAD1 (rabbit,1:1000, [54] with electrophoresis under non-denaturing conditions) .

Data availability.

All sequence datasets are available at Array Express with the accession number E-MTAB-4564.

Ethical approval

The research did not involve animals or humans and did not require ethical approval.

Acknowledgements

For antibodies we thank Keith Matthews (Edinburgh University, Scotland; PIP39 and PAD1), Isabel Roditi (University of Bern, Switzerland, BARP); Shula Michaeli (Bar-Ilan University, Israel, anti-HNRNPH), Paul MIchels (anti-PGK, Edinburgh University, Scotland) and Minu Chauduri (Meharry Medical College, Nashville, Tennessee, USA, AOX). The anti-VSG117 antibody was either from Peter Overath (Max-Planck Institut, Tübingen, Germany) or George Cross (Rockefeller University, USA). We also thank Keith, Isabel and Shula for useful discussions. All RNA-Seq libraries were prepared by David Ibberson at Bioquant, Heidelberg, and sequenced at EMBL. We thank Claudia Helbig and Ute Leibfried for technical assistance and Kevin Leiss for counting the numbers of elements in all 3'-UTRs.

References

1. Horn D (2014) Antigenic variation in African trypanosomes. Mol Biochem Parasitol 195: 123-129.
2. Vassella E, Reuner B, Yutzy B, Boshart M (1997) Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. J Cell Sci 110: 2661-2671.
3. Vassella E, Acosta-Serrano A, Studer E, Lee SH, Englund PT, et al. (2001) Multiple procyclin isoforms are expressed differentially during the development of insect forms of Trypanosoma brucei. J Mol Biol 312: 597-607.
4. Rotureau B, Van Den Abbeele J (2013) Through the dark continent: African trypanosome development in the tsetse fly. Front Cell Infect Microbiol 3: 53.
5. Urwyler S, Studer E, Renggli CK, Roditi I (2007) A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of Trypanosoma brucei. Mol Microbiol 63: 218-228.
6. Peacock L, Bailey M, Carrington M, Gibson W (2013) Meiosis and Haploid Gametes in the Pathogen Trypanosoma brucei. Curr Biol.
7. Peacock L, Ferris V, Sharma R, Sunter J, Bailey M, et al. (2011) Identification of the meiotic life cycle stage of *Trypanosoma brucei* in the tsetse fly. Proc Natl Acad Sci U S A 108: 3671-3676.

8. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, et al. (2010) Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. PLoS Pathog 6: e1001037.
9. Siegel T, Hekstra D, Wang X, Dewell S, Cross G (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Res 38: 4946-4957.
10. Fadda A, Ryten M, Droll D, Rojas F, Färber V, et al. (2014) Transcriptome-wide analysis of mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. Mol Microbiol 94: 307-326.
11. Jensen BC, Ramasamy G, Vasconcelos EJ, Ingolia NT, Myler PJ, et al. (2014) Extensive stage-regulation of translation revealed by ribosome profiling of *Trypanosoma brucei*. BMC Genomics 15: 911.
12. Vasquez JJ, Hon CC, Vanselow JT, Schlosser A, Siegel TN (2014) Comparative ribosome profiling reveals extensive translational complexity in different *Trypanosoma brucei* life cycle stages. Nucleic Acids Res 42: 3623-3637.
13. Gunasekera K, Wuthrich D, Braga-Lagache S, Heller M, Ochsenreiter T (2012) Proteome remodelling during development from blood to insect-form Trypanosoma brucei quantified by SILAC and mass spectrometry. BMC Genomics 13: 556.
14. Urbaniak MD, Guther ML, Ferguson MA (2012) Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. PLoS One 7: e36619.
15. Domingo-Sananes MR, Szoor B, Ferguson MA, Urbaniak MD, Matthews KR (2015) Molecular control of irreversible bistability during trypanosome developmental commitment. J Cell Biol 211: 455-468.
16. Dejung M, Subota I, Bucerius F, Dindar G, Freiwald A, et al. (2016) Quantitative proteomics uncovers novel factors involved in developmental dfferentiation of *Trypanosoma brucei*. PLoS Pathog 12: e1005439.
17. Michaeli S (2011) Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. Future Microbiol 6: 459-474.
18. Clayton C, Michaeli S (2011) 3' processing in protists. Wiley Interdiscip Rev RNA 2: 247-255.
19. Clayton C (2013) The regulation of trypanosome gene expression by RNA-binding proteins. PLoS Pathog 9: e1003680.
20. Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, et al. (2013) The hnRNP F/H homologue of *Trypanosoma brucei* is differentially expressed in the two life cycle stages of the parasite and regulates splicing and mRNA stability. Nucleic Acids Res 41: 6577-6594.
21. Gupta SK, Chikne V, Eliaz D, Tkacz ID, Naboishchikov I, et al. (2014) Two splicing factors carrying serine-arginine motifs, TSR1 and TSR1IP, regulate splicing, mRNA stability, and rRNA processing in Trypanosoma brucei. RNA Biol 11.
22. Antwi E, Haanstra J, Ramasamy G, Jensen B, Droll D, et al. (2016) Integrative analysis of the *Trypanosoma brucei* gene expression cascade predicts differential regulation of mRNA processing and unusual control of ribosomal protein expression. BMC Genomics 17: 306.
23. Queiroz R, Benz C, Fellenberg K, V S, Hoheisel J, et al. (2009) Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. BMC Genomics 10: 495.
24. Archer SK, van Luu D, de Queiroz R, Brems S, Clayton CE (2009) *Trypanosoma brucei* PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. PLoS Pathog 5: e1000565.

25. Droll D, Minia I, Fadda A, Singh A, Stewart M, et al. (2013) Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. PLoS Pathog 9: e1003286.
26. Kolev NG, Ullu E, Tschudi C (2014) The emerging role of RNA-binding proteins in the life cycle of *Trypanosoma brucei*. Cell Microbiol 16: 482-489.
27. Janzen CJ, van Deursen, F., Shi H, Cross GAM, Matthews KR, et al. (2006) Expression site silencing and life-cycle progression appear normal in Argonaute1-deficient *Trypanosoma brucei* Mol Biochem Parasitol 14: 102-107.
28. Tschudi C, Shi H, Franklin J, Ullu E (2012) Small interfering RNA-producing loci in the ancient parasitic eukaryote *Trypanosoma brucei*. BMC Genomics 13: 427.
29. Gunzl A, Bruderer T, Laufer G, Schimanski B, Tu LC, et al. (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei.* Eukaryot Cell 2: 542-551.
30. Biebinger S, Rettenmaier S, Flaspohler J, Hartmann C, Peña-Diaz J, et al. (1996) The PARP promoter of *Trypanosoma brucei* is developmentally regulated in a chromosomal context. Nucleic Acids Res 24: 1202-1211.
31. Furger A, Schürch N, Kurath U, Roditi I (1997) Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. Mol Cell Biol 17: 4372-4380.
32. Hehl A, Vassella E, Braun R, Roditi I (1994) A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in *Trypanosoma brucei*. Proc Natl Acad Sci USA 91: 370-374.
33. Schürch N, Furger A, Kurath U, Roditi I (1997) Contribution of the procyclin 3' untranslated region and coding region to the regulation of expression in bloodstream forms of *Trypanosoma brucei*. Mol Biochem Parasit 89: 109-121.
34. Hotz H-R, Hartmann C, Huober K, Hug M, Clayton CE (1997) Mechanisms of developmental regulation in *Trypanosoma brucei*: A polypyrimidine tract in the 3'-untranslated region of a trypanosome surface protein mRNA affects RNA abundance and translation. Nucleic Acids Res 25: 3017-3025.
35. Hotz H-R, Biebinger S, Flaspohler J, Clayton CE (1998) *PARP* gene expression: regulation at many levels. Mol Biochem Parasitol 91: 131-143.
36. Engstler M, Boshart M (2004) Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. Genes Dev 18: 2798-2811.
37. Wurst M, Selinger B, Jha B, Klein C, Queiroz R, et al. (2012) Expression of the RNA Recognition Motif protein RBP10 promotes a bloodstream-form transcript pattern in *Trypanosoma brucei*. Mol Microbiol 83: 1048-1063.
38. Telleria EL, Benoit JB, Zhao X, Savage AF, Regmi S, et al. (2014) Insights into the trypanosome-host interactions revealed through transcriptomic analysis of parasitized tsetse fly salivary glands. PLoS Negl Trop Dis 8: e2649.
39. Erben E, Fadda A, Lueong S, Hoheisel J, Clayton C (2014) Genome-wide discovery of post-transcriptional regulators in *Trypanosoma brucei*. PLoS Pathogens 10: e1004178.
40. Lueong S, Merce C, Fischer B, Hoheisel J, Erben E (2016) Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. Mol MIcrobiol 100: 457-471.
41. Zinoviev A, Leger M, Wagner G, Shapira M (2011) A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. Nucleic Acids Res 39: 8404-8415.
42. Alsford S, Turner D, Obado S, Sanchez-Flores A, Glover L, et al. (2011) High throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res 21: 915-924.

894 43. Mayho M, Fenn K, Craddy P, Crosthwaite S, Matthews K (2006) Post-
895 transcriptional control of nuclear-encoded cytochrome oxidase subunits in
896 *Trypanosoma brucei*: evidence for genome-wide conservation of life-cycle
897 stage-specific regulatory elements. Nucleic Acids Res 34: 5312-5324.
898 44. Arieti F, Gabus C, Tambalo M, Huet T, Round A, et al. (2014) The crystal
899 structure of the Split End protein SHARP adds a new layer of complexity to
900 proteins containing RNA recognition motifs. Nucleic Acids Res 42: 6742-
901 6752.
902 45. Morgan CE, Meagher JL, Levengood JD, Delproposto J, Rollins C, et al. (2015)
903 The First Crystal Structure of the UP1 Domain of hnRNP A1 Bound to RNA
904 Reveals a New Look for an Old RNA Binding Protein. J Mol Biol 427: 3241-
905 3257.
906 46. Wang X, Tanaka Hall T (2001) Structural basis for recognition of AU-rich element
907 RNA by the HuD protein. Nat Struct Biol 8: 141-145.
908 47. Deo R, Bonanno J, Sonenberg N, Burley S (1999) Recognition of polyadenylate
909 RNA by the poly(A)-binding protein. Cell 98: 835-845.
910 48. Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, et al. (1999) Structural basis
911 for recognition of the tra mRNA precursor by the Sex-lethal protein. Nature
912 398: 579-585.
913 49. Drozdz M, Clayton CE (1999) Structure of a regulatory 3'-untranslated region
914 from *Trypanosoma brucei*. RNA 5: 1632-1644.
915 50. Misset O, Opperdoes FR (1987) The phosphoglycerate kinases from
916 *Trypanosoma brucei*. Eur J Biochem 162: 493-500.
917 51. Gibson WC, Swinkels BW, Borst P (1988) Post-transcriptional control of the
918 differential expression of phosphoglycerate kinase genes in *Trypanosoma*
919 *brucei*. J Mol Biol 201: 315-325.
920 52. Blattner J, Clayton CE (1995) The 3'-untranslated regions from the *Trypanosoma*
921 *brucei* phosphoglycerate kinase genes mediate developmental regulation.
922 Gene 162: 153-156.
923 53. Quijada L, Hartmann C, Guerra-Giraldez C, Drozdz M, Irmer H, et al. (2002)
924 Expression of the human RNA-binding protein HuR in *Trypanosoma brucei*
925 induces differentiation-related changes in the abundance of developmentally-
926 regulated mRNAs. Nucleic Acids Res 30: 1-11.
927 54. Dean S, Marchetti R, Kirk K, Matthews K (2009) A surface transporter family
928 conveys the trypanosome differentiation signal. Nature 459: 213-217.
929 55. Szoor B, Ruberto I, Burchmore R, Matthews KR (2010) A novel phosphatase
930 cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal
931 signaling pathway. Genes Dev 24: 1306-1316.
932 56. Benz C, Mulindwa J, Ouna B, Clayton C (2011) The *Trypanosoma brucei* zinc
933 finger protein ZC3H18 is involved in differentiation. Mol Biochem Parasitol
934 177: 148-151.
935 57. Kershaw CJ, Costello JL, Talavera D, Rowe W, Castelli LM, et al. (2015)
936 Integrated multi-omics analyses reveal the pleiotropic nature of the control of
937 gene expression by Puf3p. Sci Rep 5: 15518.
938 58. Ling A, Trotter J, Hendriks E (2011) A zinc finger protein, TbZC3H20, stabilises
939 two developmentally regulated mRNAs in trypanosomes. J Biol Chem 286:
940 20152-20162.
941 59. Jones NG, Thomas EB, Brown E, Dickens NJ, Hammarton TC, et al. (2014)
942 Regulators of Trypanosoma brucei cell cycle progression and differentiation
943 identified using a kinome-wide RNAi screen. PLoS Pathog 10: e1003886.
944 60. Vassella E, Probst M, Schneider A, Studer E, Renggli C, et al. (2004) Expression
945 of a major surface protein of *Trypanosoma brucei* insect forms is controlled by
946 the activity of mitochondrial enzymes. Mol Biol Cell 15: 3986-3993.

947 61. Kolev NG, Ramey-Butler K, Cross GA, Ullu E, Tschudi C (2012) Developmental
948 progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding
949 protein. Science 338: 1352-1353.
950 62. Lott K, Mukhopadhyay S, Li J, Wang J, Yao J, et al. (2015) Arginine methylation
951 of DRBD18 differentially impacts its opposing effects on the trypanosome
952 transcriptome. Nucleic Acids Res 43: 5501-5523.
953 63. Minia I, Merce C, Terrao M, Clayton C (2016) Many mRNAs that increase after
954 heat shock of procyclic-form *Trypanosoma brucei* are also induced during
955 differentiation to mammalian-infective forms. submitted.
956 64. Wang H, Li X, Gao S, Sun X, Fang H (2015) Transdifferentiation via transcription
957 factors or microRNAs: Current status and perspective. Differentiation 90: 69-
958 76.
959 65. van der Meulen T, Huising MO (2015) Role of transcription factors in the
960 transdifferentiation of pancreatic islet cells. J Mol Endocrinol 54: R103-117.
961 66. Gao T, McKenna B, Li C, Reichert M, Nguyen J, et al. (2014) Pdx1 maintains
962 beta cell identity and function by repressing an alpha cell program. Cell Metab
963 19: 259-271.
964 67. Alibu VP, Storm L, Haile S, Clayton C, Horn D (2004) A doubly inducible system
965 for RNA interference and rapid RNAi plasmid construction in *Trypanosoma*
966 *brucei*. Mol Biochem Parasitol 139: 75-82.
967 68. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, et al. (2013)
968 Comparative analysis of RNA sequencing methods for degraded or low-input
969 samples. Nat Methods 10: 623-629.
970 69. Estévez A, Kempf T, Clayton CE (2001) The exosome of *Trypanosoma brucei*.
971 EMBO J 20: 3831-3839.
972 70. Singh A, Minia I, Droll D, Fadda A, Clayton C, et al. (2014) Trypanosome MKT1
973 and the RNA-binding protein ZC3H11: interactions and potential roles in post-
974 transcriptional regulatory networks. Nucleic Acids Res 42: 4652-4668.
975 71. Minia I, Clayton C (2016) Regulating a post-transcriptional regulator: protein
976 phosphorylation, degradation and translational blockage in control of the
977 trypanosome stress-response RNA-binding protein ZC3H11. PLoS
978 Pathogens 12: e1005514.
979 72. Martin M (2011) Cutadapt removes adapter sequences from high-throughput
980 sequencing reads. EMBnetjournal 17: 10-12.
981 73. Langmead B, Trapnell C, Pop M, Salzberg S (2009) Ultrafast and memory-
982 efficient alignment of short DNA sequences to the human genome. Genome
983 Biol 10: R25.
984 74. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence
985 Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.
986 75. Anders S, Huber W (2010) Differential expression analysis for sequence count
987 data. Genome Biol 11: R106.
988 76. Bailey T (2011) DREME: Motif discovery in transcription factor ChIP-seq data.
989 Bioinformatics 27: 1653-1659.
990 77. Kolev N, Franklin J, Carmi S, Shi H, Michaeli S, et al. (2010) The transcriptome of
991 the human pathogen *Trypanosoma brucei* at single-nucleotide resolution.
992 PLoS Pathog 6: e1001090.
993 78. Zomer AW, Allert S, Chevalier N, Callens M, Opperdoes FR, et al. (1998)
994 Purification and characterisation of the phosphoglycerate kinase isoenzymes
995 of *Trypanosoma brucei* expressed in *Escherichia coli*
996 . Biochim Biophys Acta 1386: 179-188.
997 79. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, et al. (2012) Fiji:
998 an open-source platform for biological-image analysis. Nat Methods 9: 676-
999 682.

80. Fritz M, Vanselow J, Sauer N, Lamer S, Goos C, et al. (2015) Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. Nucl Acids Res 43: 8013-8032.

81. Mony BM, Macgregor P, Ivens A, Rojas F, Cowton A, et al. (2013) Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*. Nature 505: 681-685.

Supporting information

S1 Text

3'-untranslated regions of mRNAs that were at least 3x enriched in both RBP10 pull-downs.

S2 Text

3'-untranslated regions of mRNAs that were less than 0.7x enriched in both RBP10 pull-downs.

S3 Text

3'-untranslated regions of mRNAs that were at least 3x enriched on average in RBP10 pull-downs, according to DESeq (Padj <8 E-5), but one replicate was less than 3x enriched.

S1 Table

Proteins that co-purified with TAP-tagged RBP10.

Endogenously TAP-tagged RBP10 was purified three times in the presence or absence of RNase. The controls were three independent purifications of GFP-TAP

Sheet 1: Proteins enriched in the purifications. Peptide counts and ratios are shown. The highest peptide count from the GFP-TAP replicates was used as a baseline. The proteins shown in this table were present as at least 2 peptides, present in at least two similar purifications, and were on average at least 2-fold enriched relative to GFP-TAP. Functions of interest are highlighted in different colours. "Average PGKB =tet/-tet" are tethering screen results [39]; a value above 1.5 suggests that the protein can suppress expression. "in starvation granules?" is 1 if positive and 0 if negative [80]. The number of peptides found by purification of MKT1-TAP are also indicated [70].

Sheet 2: All raw data.

S2 Table

Results of a yeast 20-hybrid screen with a random fragment library.

The library [70] was screened with pGAD-RBP10 and results evaluated by sequencing.

Sheet 1: Raw data for pGAD-RBP10. All reads that started from position -36 onwards (relative to the ATG) are included.

Sheet 2: Raw data for pGAD-T7 (unselected background, filtered for at least 4 reads)

Sheet 3: Results for pGAD-RBP10, filtered to exclude loci with less than 10 reads, and open reading frames with less than 2 independent fragments. Numerous cysteine protease fragments have been removed.

Sheet 4: Possible interaction partners. The ratio to the maximum background count is at least 4. Proteins that also scored as interacting with MKT1 or CFB1 [70] (indicating less specificity) are listed separately. The "regulator" column shows results from the tethering screen [39].

S3 Table

Effect of RBP10 depletion on gene expression in bloodstream forms.

RNAi was induced for 15h and mRNA was collected for RNASeq from total cells, the free and monosomal fractions, and the polysomal fractions. Data are for two independent experiments which are labelled A and B.

Throughout the Tables, the colour code for ratios is as follows: Red is >2, orange is 1.5-2.0, green is 0.667 - 0.5 and blue is <0.5. Shading of columns does not follow a particular colour code but is designed to make the tables easier to navigate. Many numbers have been rounded to 2 or 3 points beyond the decimal in order to decrease the sizes of the files. Ribosome profiling results (Riboprof) are from [11], and developmental regulation data for total RNA are from [10].

Sheet 1: Raw read counts for all unique genes, chosen according to [9] with some later additions.

Sheet 2: Columns E - P Results were converted to reads per million reads (RPM). For each experiment, we then corrected the RPM for the monosomes and polysomes according to the fraction of the total mRNA that was in each fraction (sheet 7 and S3 Fig, numbers not included in the Table). We then calculated the percentage of each mRNA that was on polysomes (columns Q-T). Column U contains the average of

ratio between RNAi and control, calculated for the total input RNA - (K+L) / (E+F). Column V is the same ratio for the fraction of RNA not on polysomes, and column W is the ratio for the polysomal fractions (S+T) / (Q+R). To find the lowest possible estimate of the ratio in W, we took the lowest value for the proportion in polysomes after *rbp10* RNAi, and divided by the highest value from wild-type. From this column (X) we could find mRNAs that reliably move *into* polysomes after RNAi. Column Y is similar but shows the *highest* estimate for the ratio; this can be used to find mRNAs that reliably move *away* from polysomes after RNAi. Since there were only 3 genes in this category (column Y) they were not considered further.

Sheet 3: mRNAs that move onto polysomes after *rbp10* RNAi. Same as sheet 2, except that only genes for which column W is more than 1.5, and column X is more than 1.2, are included. (This means that the average increase is at least 1.5x and the minimum is 1.2x.) Columns Y and Z are ribosome profiling results from [11], giving the ribosome density ratios for bloodstream forms (BS) divided by procyclic forms (PC). Slender bloodstream forms from blood are slBS and cultured bloodstream forms are cBS. Column AB is results from S5 Table, sheet 2 (binding of the RNA to RBP10, ratio of bound to unbound) and column AC is the adjusted P-values for the results in column AB.

Sheet 4: DeSeq analysis of the data in Sheet 1.

Sheet 5: Genes showing at least a 1.5-fold increase either in the amount of total RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-value threshold 0.05). Column AA is the RNA-binding ratio from S5 Table, sheet 2,

Sheet 6: As sheet 5, except that these are genes showing at least a 1.5-fold decrease either in the amount of total RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-value threshold 0.05).

Sheet 7: Quantitation of spliced leader signals for the free/monosome and polysome fractions used for RNA-Seq

Sheet 8: Statistical analyses to compare the proteomes of differentiating trypanosomes [16] with the effects of RBP10 RNAi on RNA levels.

S4 Table

Effect of RBP10 expression on gene expression in procyclic forms.

Expression of RBP10 was induced with tetracycline addition for 6 hours, and RNA samples prepared from duplicates as in Supplementary Table S3. Controls are the same cells without tetracycline.

Sheet 1: Columns D-O are raw reads. Columns P-S are corresponding RPM values for input samples.

Sheet 2: Proportions in free/monosome and polysome fractions were calculated as described for Supplementary Table S3, sheet 2. Columns Q-X: Effects on both polysomes and free/monosome fractions were also calculated in a similar way.

Sheet 3: As sheet 2, but only genes showing regulated translation according to the following criteria: (a) movement to the free/monosome fraction from the polysomes. This is the top section (rows 2-320). The smallest value for the proportion in free/monosomes with RBP10 expression was divided by the largest value calculated value for the proportion in free/monosomes without RBP10 induction. If this number was greater than 1.5, the gene was included. In addition, we divided the largest number for the proportion in polysomes plus RBP10 expression by the smallest value for the proportion in polysomes without RBP10 induction. If this number was less than 0.667 the gene was also included. Genes that satisfied both criteria are in bold. Additional columns are as in Table S3. Rows 322-326 are the four RNAs showing movement onto polysomes.

Sheet 4: DeSeq analysis of the results in Sheet 1.

Sheet 5: Genes showing at least a 1.5-fold increase either in the amount of total RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-value threshold 0.05).

Sheet 6: As sheet 5, except that these are genes showing at least a 1.5-fold decrease either in the amount of total RNA, or the amount of RNA in polysomes, or both, according to DeSeq (adjusted P-value threshold 0.05). Columns V-Z are results for salivary gland trypanosomes, calculated from the raw data in [38]; statistical analysis was not possible due to the absence of replicas.

Sheet 7: Quantitation of spliced leader signals for the free/monosome and polysome fractions used for RNASeq.

S5 Table

RNAs bound to tagged RBP10.

Results for two independent experiments, A and B, are shown.

Sheet 1: Raw read counts for all samples, calculation of reads per million reads (RPM), and ratios of bound/unbound for each experiment.

Sheet 2: DeSeq analysis of the raw data in sheet 1

Sheet 3: Set of mRNAs that were at least 3-fold bound according to DeSeq. This is a slightly more relaxed stringency than that used in sheet 5. The largest P-value was 7E-5. H-J: Results for developmental regulation [10]; K 0- calculated from [38]; M-T: results from Supplementary Tables S3 and S4; U, W, Z: [13]; V: [14]; X, Y, AA: [15]; AB: [16]; AC: [81]; AD: [42].

Sheet 4: Statistics used to create the Venn diagrams in Figure 2.

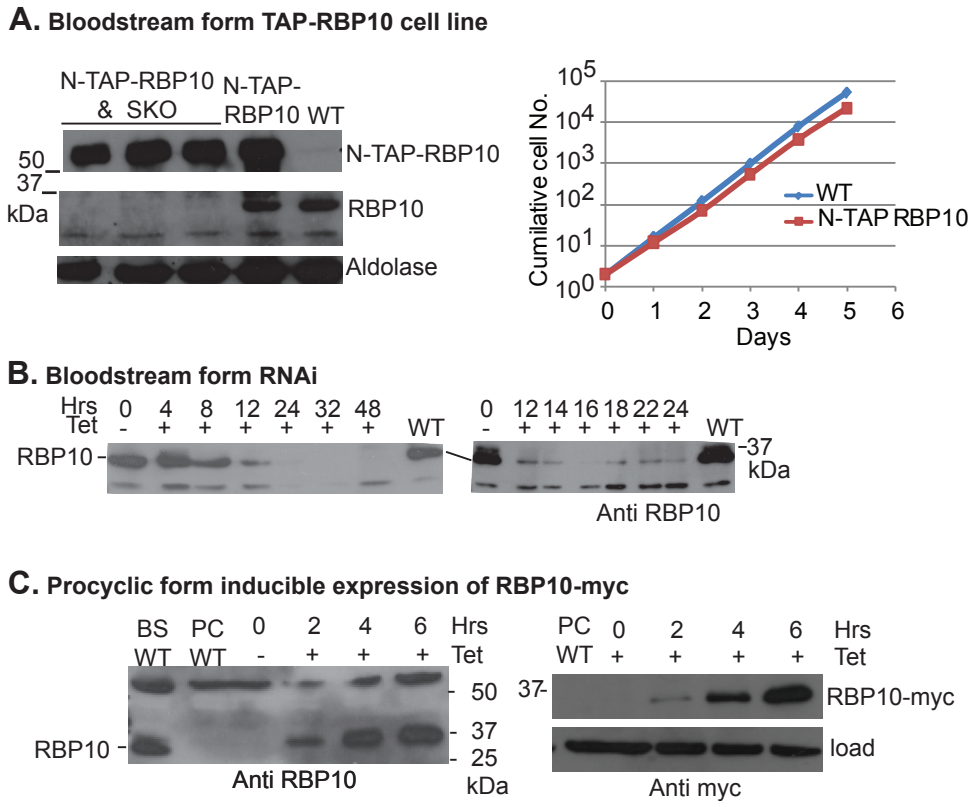
Sheet 5: Set of bound RNAs used for the motif search. Filters were (a) at least 10 reads in all samples and (b) a bound/unbound ratio of >3 in both experiments. The 3'-UTR lengths are those annotated in TritrypDB, except where noted. Only annotated 3'-UTRs were used for the initial motif search.

S6 Table

Plasmids and oligonucleotides

S1 Fig

Alignment of RBP10 from different kinetoplastids



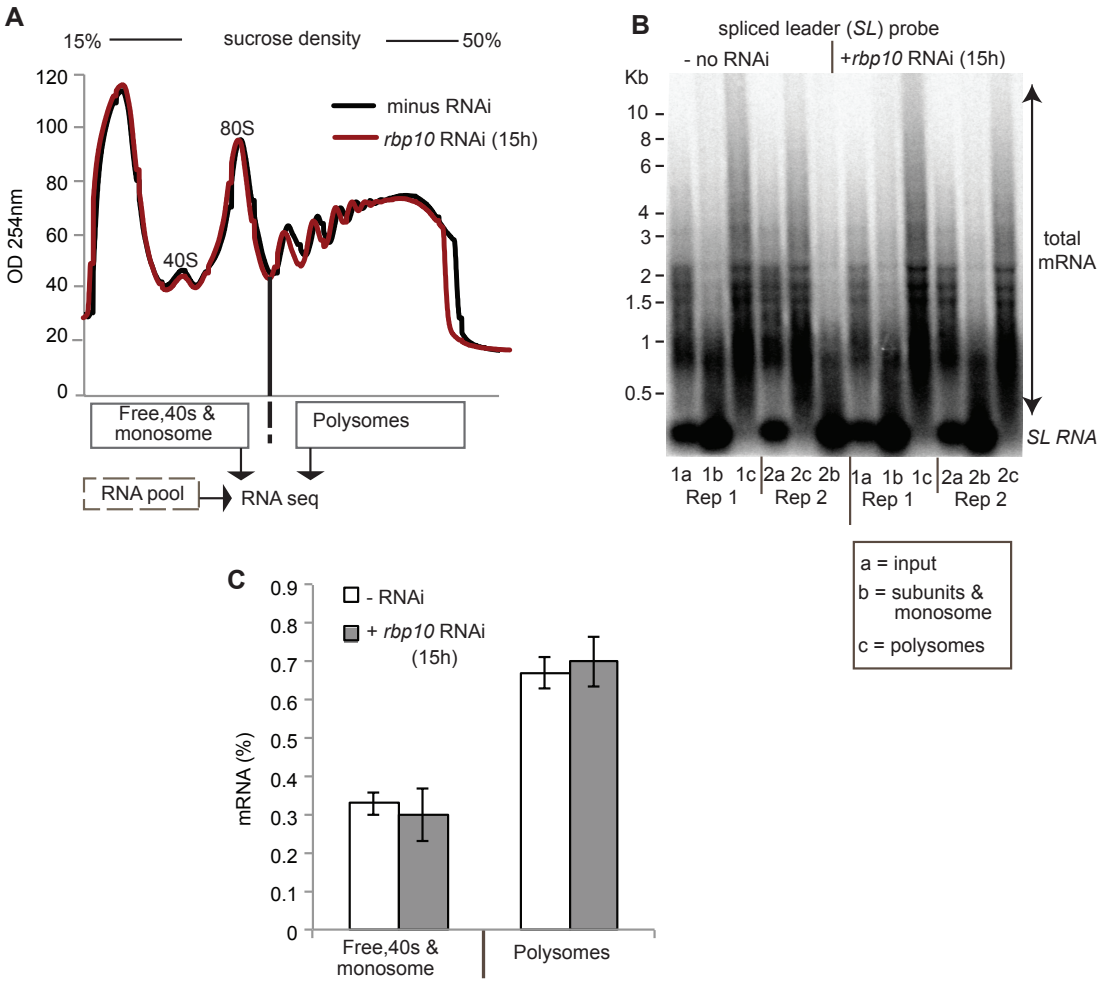
S2 Fig

Cell line characterisation and RBP10 expression

A) Characterisation of bloodstream-form trypanosomes with a sequence encoding an N-terminal tandem affinity purification tag integrated in frame with the *RBP10* coding region. The left hand panel is a western blot showing the presence or absence of TAP-RBP10 and native RBP10, with genotypes above the lanes. The right-hand panel is a cumulative growth plot for the cell line used in the interactome experiments.

B) RNAi cell line: Western blots showing the time course of RBP10 decrease after tetracycline addition

C) Western blots showing the time course of expression of RBP10-myc in Lister 427 procyclic forms after tetracycline addition.



1161

1162 **S3 Fig**

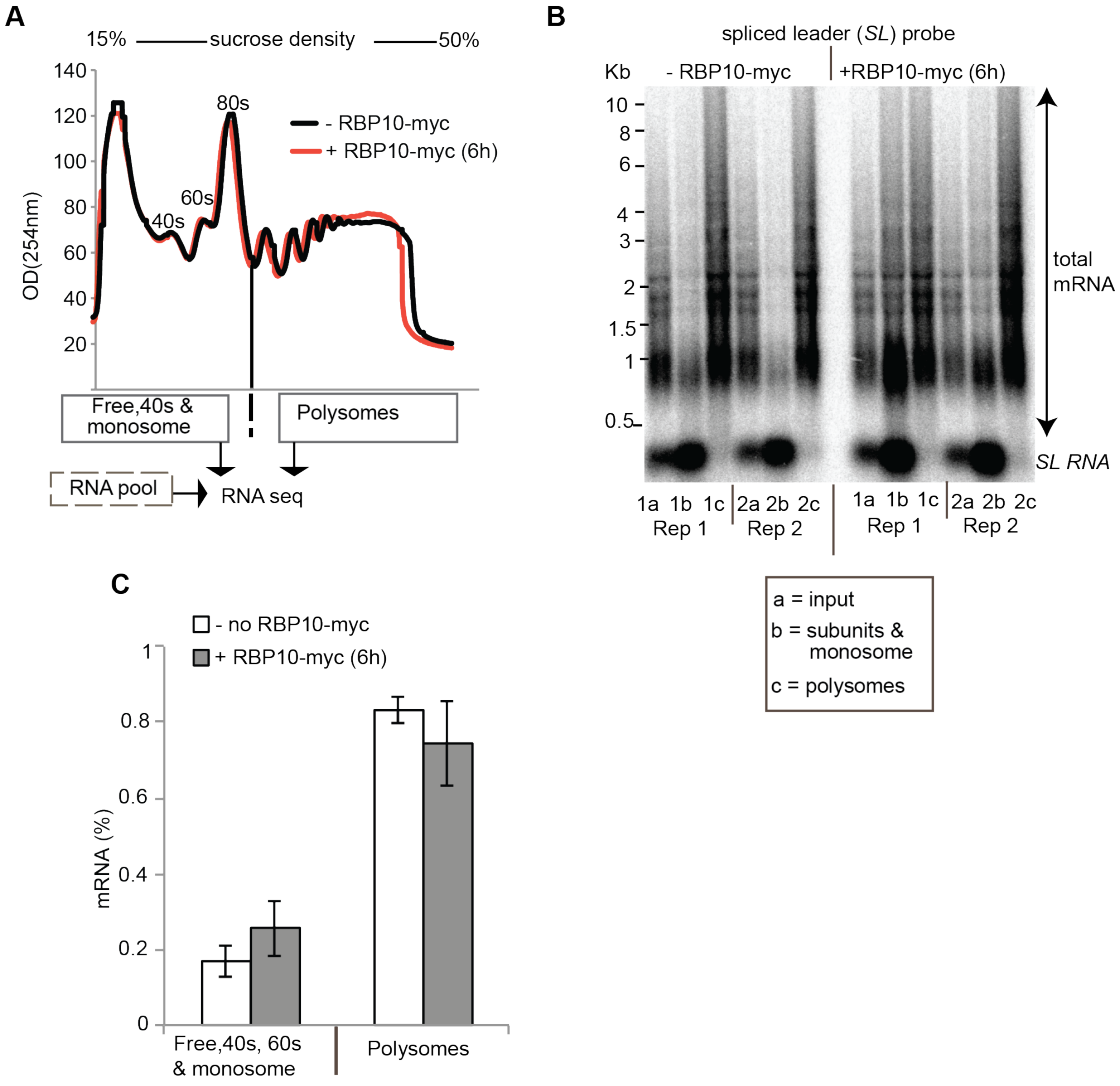
1163 RNAi targeting *RBP10* in bloodstream forms: samples used for RNASeq

1164 A. Typical Sucrose gradient profiles of samples

1165 B. Total mRNA in the sucrose gradient fractions, detected using the spliced leader
1166 probe

1167 C. Average quantitation of the spliced leader signal

1168



1169

1170 **S4 Fig**

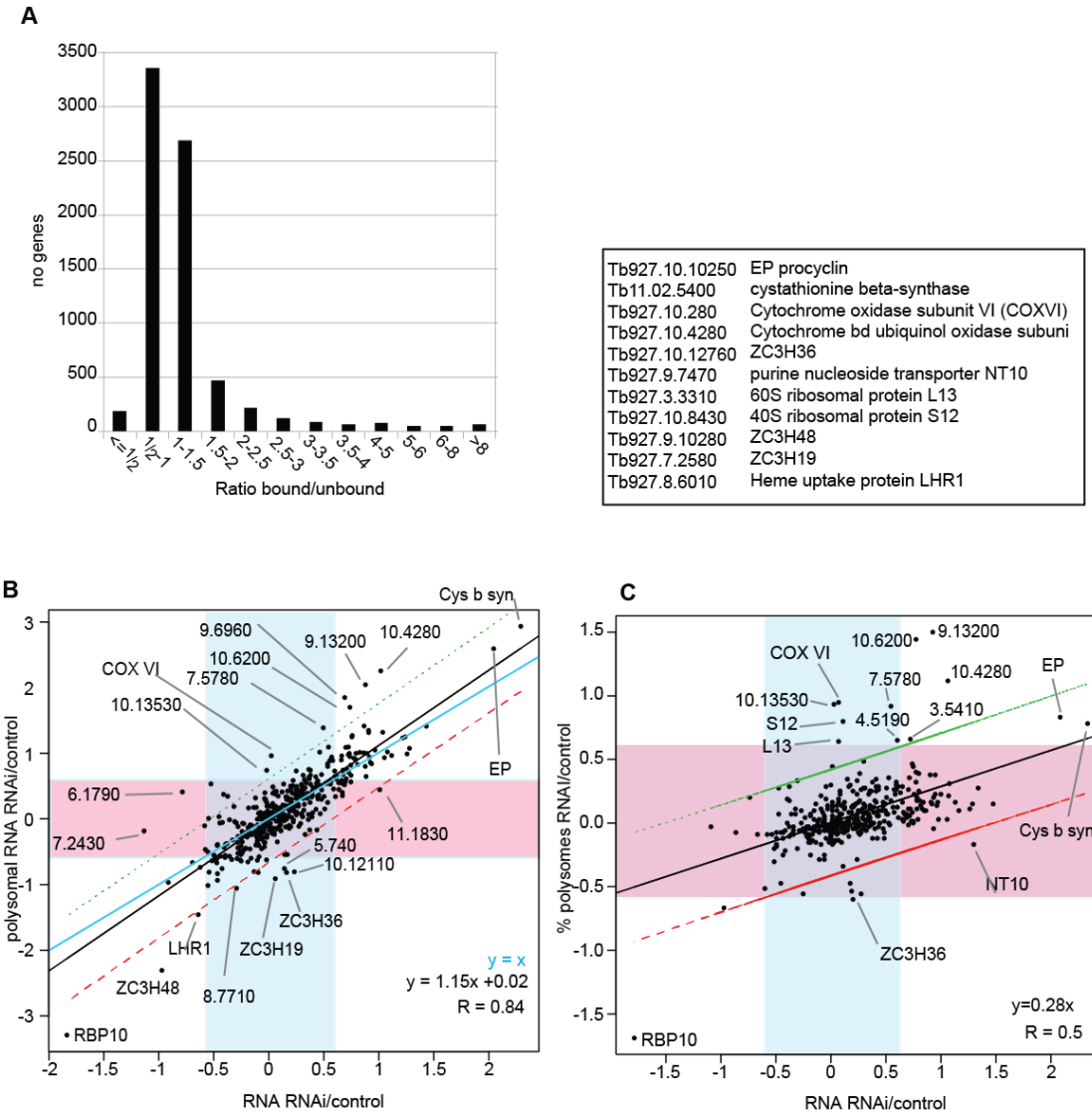
1171 Expression of RBP10 in procyclic forms: samples used for RNASeq

1172 A. Typical Sucrose gradient profiles of samples

1173 B. Total mRNA in the sucrose gradient fractions, detected using the spliced leader
1174 probe,

1175 C. Average quantitation of the spliced leader signal

1176



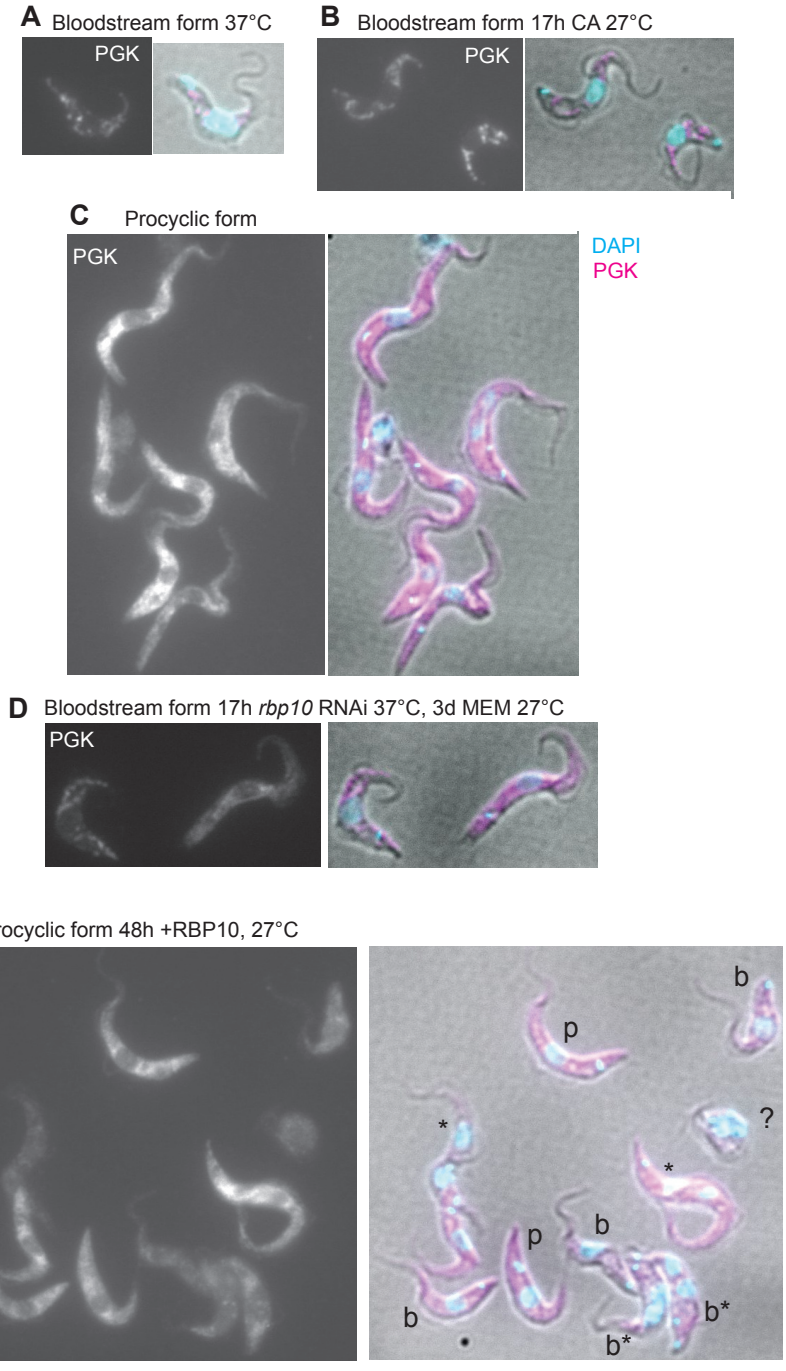
S5 Fig

Effects of RNAi and analysis of RBP10-bound RNAs.

A. Analysis of binding of individual mRNAs to RBP10. The classes are ratios of bound/unbound; the number of different open reading frames in each class is on the y axis.

B. Scatter plot comparing the effect of *rbp10* RNAi on total RNA (x axis) with the effect on polysomal RNA (y axis) for mRNAs that were at least 3x enriched according to DeSeq; all P-values were less than 8E-5. The black line is the regression line and the red and green lines show the 95% confidence limits for the data. The blue shadow encloses total mRNAs that were less than 1.5x affected, and the pink shadow encloses polysomal RNAs that were less than 1.5x affected. The cyan line is perfect correlation. The box beneath the graph lists relevant TritrypDB accession numbers. The gene numbers on the plot are also accession numbers, with "Tb927." removed.

C. As (B), but here the y axis shows the effect of RNAi on the percentage of the mRNA in polysomes.



S6 Fig

Relocation of phosphoglycerate kinase during differentiation.

Cells were fixed with fomaldehyde, pernmeabilized with triton x-100, and stained using a polyclonal antibody to phosphglycerate kinase (PGK). The grey-scale panels show PGK alone and the differential interference contrast panels show DNA in cyan and PGK in magenta.

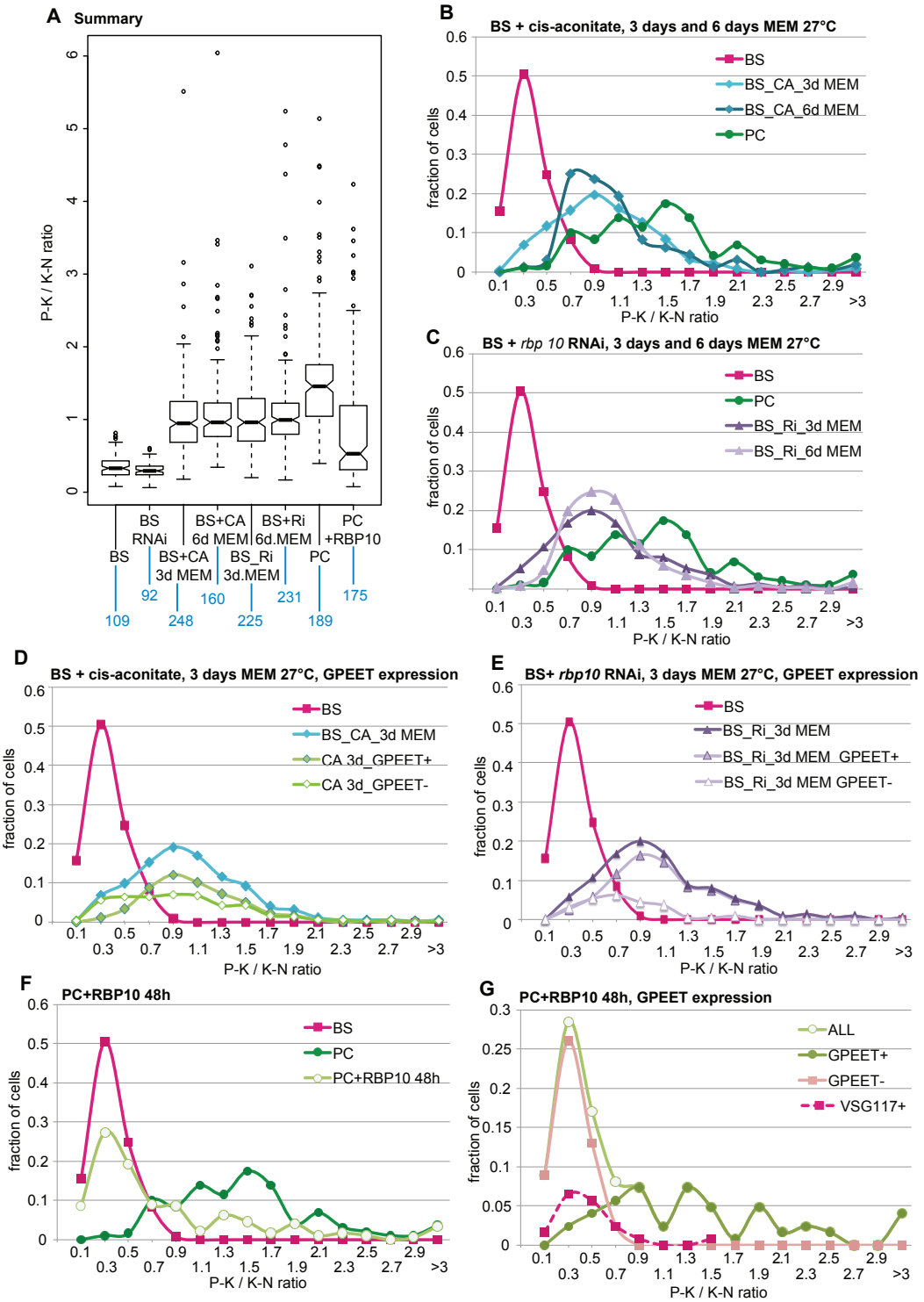
A. Bloodstream forms

B. Bloodstream forms after incubation with cis aconitate for 17h at 27°C.

C. Procyclic forms.

D. Bloodstream forms with 17h *rbp10* RNAi followed by culture for 3 days under procyclic-form culture conditions. The selection shows on procyclic-like trypanosome (left) and one which still has bloodstream-form morphology (right).

E. Procyclic forms after 2 days induction of expression of RBP10-myc. Cells with terminal kinetoplasts are labelled "b" and cells with more procyclic morphology are labelled "p". Cells that appear to be dividing are indicated with asterisks. A very abnormal cell is indicated with "?". Cells if unclear status are not labelled.



S7 Fig

Morphological analysis of differentiation.

Cells were treated as indicated, fixed, and stained for DNA (cyan), EP-procyclin, phopho-GPEET procyclin, or VSG. (magenta). The posterior-kinetoplast and kinetoplast-nucleus distances were measured in Image J and the ratios calculated.

A. Summary of all results. This is the same as Fig 5H, but all of the outliers are included, and the numbers of cells analysed are below each lane. BS - normal bloodstream forms; PC- normal procyclic forms; CA - 17h cis-aconitate pretreatment; RNAi or Ri; 17h *rbp10* induction; +RBP10 - induced expression or RBP10 for 2 days.

B. Cell distributions after cis-aconitate-induced differentiation. The x axis shows the P-K/K-N ration, and y-axis shows the percentage of cells with that ratio. Controls are bloodstream forms (magenta) and established procyclic forms (green). The other lines are for 3 and 6 days after transfer to procyclic medium (MEM).

C. As B, but with transfer after 17h *rbp10* RNAi.

D. Phospho-GPEET expression 3 days after cis-aconitate-stimulated differentiation. The cells indicated with the cyan line (BS_CA_3d MEM) were subdivided into GPEET positive and GPEET negative.

E. Phospho-GPEET expression 3 days after RNAi-stimulated differentiation; other details as in D.

F. Cell distributions in procyclic forms after induced expression of RBP10.

G. As (F), but also showing staining with anti-phospho-GPEET and anti-VSG117. Cells wit no GPEET have bloodstream-form kinetoplast positions, and a subset of them cross-reacts with anti-VSG117 antibodies. The remainder may express VSGs that do not cross react with the anti-117 antibody.